DOCUMENT RESUME

ED 150 023 SE 023 940

AUTHOR	Helgeson, Stanley L., Ed.; Blosser, Patricia E., Ed.
TITLE	Investigations in Science Education, Vol. 2, No. 4. Expanded Abstracts and Critical Analyses of Recent Research.
INSTITUTION	Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE	76
NOTE	66p.
AVAILABLE FROM	Information Reference Center (ERIC/IRC), The Ohio State University, 1200 Chambers Rd., 3rd Floor, Columbus, Ohio 43212 (Subscription, $6.00, $1.75 ea.)
EDRS PRICE	MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS	*Abstracts; Biology; Chemistry; *Educational Research; *Instruction; Physics; Research; Science Education; *Student Behavior; *Teacher Behavior
IDENTIFIERS	*Research Reports

ABSTRACT

Presented are analytical abstracts, prepared by science educators, of fourteen research reports. The abstracts are grouped in three clusters: teacher characteristics and behaviors, student characteristics and behaviors, and instruction. Each abstract includes bibliographical data, research design and procedure, purpose, research rationale, and an abstractor's analysis of the research. (BB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED150023

INVESTIGATIONS IN SCIENCE EDUCATION

Editor

Stanley L. Helgeson
The Ohio State University

Associate Editor

Patricia E. Blosser
The Ohio State University

Advisory Board

Mary Budd Rowe
University of Florida

Robert G. Bridgham
Michigan State University

Kenneth G. Jacknicke
University of Alberta

Vincent N. Lunetta
University of Iowa

Robert Collagen
Towson State College

David P. Butts
University of Georgia

National Association for Research in Science Teaching

ERIC Center for Science, Mathematics
and Environmental Education

Published Quarterly by

The Center for Science and Mathematics,Education
College of Education
The Ohio State University
1945 North High Street
Columbus, Ohio 43210

Subscription Price: $6.00 per year. Single Copy Price: $1.75.
Add 25¢ for Canadian mailings and 50¢ for foreign mailings.

SE 023 940

Volume 2, Number 4, 1976

NOTES from the Editor iii

TEACHER CHARACTERISTICS AND BEHAVIORS 1

Dietrich, Don. "Grading Practices of High School Physics Teachers: A Contributing Factor to Declining Enrollments in Physics." Science Education, 57(1):25-29, 1973.
Abstracted by RODNEY L. DORAN 3

Lawrence, Frances. "Science Teachers' Perceptions of Their Teaching Skills and Their School Conditions." Science Education, 58(4):489-496, 1974.
Abstracted by JUDY EGELSTON-DODD 8

Mayer, Victor J., John F. Disinger, and Arthur L. White. "Evaluation of an Inservice Program for Earth Science Teachers." Science Education 59(2):145-153, 1975.
Abstracted by HERBERT A. SMITH 12

Perkes, Victor A. "Relationships Between a Teacher's Background and Sensed Adequacy to Teach Elementary Science." Journal of Research in Science Teaching, 12(1):85-88, 1975.
Abstracted by EUGENE L. CHIAPPETTA 15

Sayer, Irwin, James R. Campbell, and Cyrus W. Barnes. "The Effect of College Instructors' Interaction Ratios on Cognitive Development." Science Education, 56(4):529-537, 1972.
Abstracted by DOROTHY GABEL 19

STUDENT CHARACTERISTICS AND BEHAVIORS 23

Horn, Jerry G. "Risk-Taking in Explanation of Biological Events." Journal of Research in Science Teaching, 10(4): 341-346, 1973.
Abstracted by DAVID R. STRONCK 25

Koelsche, Charles L. and S. Lloyd Newberry. "A Study of the Relationship between Certain Variables and the Science Interest of Children." Journal of Research in Science Teaching, 8(3):237-241, 1971.
Abstracted by HANS O. ANDERSON 31

Peterson, Rita and Lawrence Lowery. "The Use of Motor Activity as an Index of Curiosity in Children." Journal of Research in Science Teaching, 9(3):193-199, 1972.
Abstracted by ANN C. HOWE 36

Simmons, Jack and William Essler. "Investigating the Attitudes Toward Science Fostered by the Process Approach Program." School Science and Mathematics, 72(7):633-636, 1972.
Abstracted by DAVID P. BUTTS 40

INSTRUCTION 43

Holliday, William G. "The Effects of Verbal and Adjunct Pictorial-Verbal Information in Science Instruction." Journal of Research in Science Teaching, 12(1):77-83, 1975.
Abstracted by ROBERT H. EVANS and RONALD D. ANDERSON 45

Howe, Ann C. and David P. Butts. "The Effect of Instruction on the Acquisition of Conservation of Volume." Journal of Research in Science Teaching, 7(4):371-375, 1970.
Abstracted by MICHAEL SZABO 49

Humphreys, Donald W. and Ronald D. Townsend. "The Effects of Teacher- and Student-Selected Activities on the Self-Image and Achievement of High School Biology Students." Science Education, 58(3):295-301, 1974.
Abstracted by GENE GENNARO 53

Kempa, R. F. and J. E. Ward. "The Effect of Different Modes of Task Orientation on Observational Attainment in Practical Chemistry." Journal of Research in Science Teaching. 12(1):69-76, 1975.
Abstracted by VINCENT N. LUNETTA 57

Sparks, Phillip D. and Laraine M. Unbehaun. "Achievement of Audio-Tutorial and Conventional Biology Students, A Comparative Study." BioScience, 21(2):574-576, 1971.
Abstracted by WILLIAM S. LaSHIER, JR. 62

NOTES . . .

from the Editor

Analyses of research reports are grouped in three clusters in this issue of I.S.E. The first cluster, TEACHER CHARACTERISTICS AND BEHAVIORS, contains five studies. The second, STUDENT CHARACTERISTICS AND BEHAVIORS, includes four reports. The last section, INSTRUCTION, includes five studies. Studies included in any given cluster do not, of course, all contain similar elements. Rather, they are grouped together because they share some common basis for comparison.

Publishable responses to the analyses and to the grouping of studies are encouraged.

Stanley L. Helgeson
Editor

Patricia E. Blosser
Associate Editor

TEACHER CHARACTERISTICS AND BEHAVIORS

Dietrich, Don. "Grading Practices of High School Physics Teachers: A Contributing Factor to Declining Enrollments in Physics." Science Education, 57(1):25-29, 1973.

Descriptors--Educational Research, *Grading, Physics, Science Education, Secondary School Science, *Student Enrollment, *Student Science Interests, Teacher Characteristics, *Teacher Influence

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Rodney L. Doran, State University of New York at Buffalo.

Purpose

This study attempted to determine if grading practices of physics teachers act as a deterrent to students enrolling in physics.

Rationale

The decrease of enrollments in high school physics in recent years has been viewed with considerable concern. Because high school physics instruction is viewed more as an important link to one's scientific literacy than solely as preprofessional preparation, this concern demands more attention. Many factors have been suggested as factors which may be related to this declining enrollment phenomenon, such as: difficulty of the course, scarcity of qualified teachers, and severity of physics teachers' grading practices. Bridgham and Welch (3) explored the relationship between grading practices and student dropouts from physics classes. Although their findings were not statistically significant, they detected trends such that they suggested further research should consider the role physics teachers' grading practices might act as a deterrent to initial enrollment in physics classes.

Research Design and Procedure

Dietrich identified schools from a midwestern state as either high enrollment (HE) or low enrollment (LE) with respect to physics enrollment. For the entire state, the percent enrollment in physics classes (computed as a proportion of 12th grade enrollment) ranged from 0 percent to 62 percent. Schools identified as HE were those with 25 percent or more of their twelfth graders enrolling in physics, while LE schools were those with 12 percent or less enrollment in physics. This arbitrary definition provided 17 HE schools and 18 LE schools with 18 teachers associated with each group of schools. Grading data were obtained from five students randomly selected from the physics class(es) of each of the teachers. The data included: (1) physics grade, (2) overall grade point average, (GPA), and (3) grade point average in science classes other than physics. Based on their grades, each student was placed into one of these categories:

I. Physics < Overall and < Other Science
II. Physics < Overall but ≥ Other Science
III. Physics ≥ Overall but < Other Science
IV. Physics ≥ Overall and ≥ Other Science

Based on the distribution of students from HE and LE schools into these categories, one is able to make comparisons as to severity of grading practices and physics enrollment.

Findings

Based on a chi-square analysis of the percent of students in the four grading categories by the enrollment variable (HE or LE), Dietrich found that there were significant differences. Most of the variation was found in categories I and IV. Sixty-two percent of teachers from HE school awarded grades in physics that were lower than students' science GPA and their overall GPA, whereas only 45 percent of the teachers from LE schools did so. Similarly, 45 percent of the LE teachers awarded physics grades that were greater than the students' overall and science GPA, while only 23 percent of the teachers from HE schools awarded such relatively high physics grades.

A grouping of categories, I with II and III with IV, allowed Dietrich to analyze physics grades earned with students' overall GPA. Inspection of these data revealed significant differences with a greater percent of teachers from HE schools (75 percent HE, 54 percent LE) awarding physics grades lower than students' overall GPA.

Grouping of categories I with III and II with IV allowed one to compare grades attained in physics classes with those obtained from other science classes. Significant differences were found with LE teachers awarding physics grades that were greater than or equal to students' "other science GPA" more frequently than do HE teachers (54 percent to 36 percent). Conversely, teachers from HE schools awarded grades in physics lower than students' "other science GPA" more frequently than did LE teachers (64 percent to 46 percent).

Interpretations

Based on the data collected and the analyses calculated, the author concluded that:

(1) Although both groups of teachers tended to be severe graders, the physics teachers in the LE schools were more severe than their counterparts in the HE schools.

(2) Comparing physics grades to overall GPA, both groups of teachers tended to award physics grades lower than the students' overall GPA, with teachers from HE schools being significantly more severe than LE physics teachers.

(3) When grades obtained in physics were compared to those earned in other science classes, the HE teachers were again more

severe in their grading practices, awarding a greater percentage of grades which were less than the students' "other science GPA", than the LE teachers.

The findings obtained in this study did not support the expectation that the severity of physics teachers' grading practices discouraged students from enrolling in physics classes. The investigator suggested that teacher classroom behavior might be a fruitful area in which to pursue this question further.

ABSTRACTOR'S ANALYSIS

The relationship between severity of grading and enrollment in high school science courses was also explored by Bridgham (2). After proposing a model of this relationship, Bridgham collected data from 27 high schools to empirically test his propositions. Consistent with Dietrich's findings, Bridgham reported that physics grades were more severe than those obtained in other non-science classes. When comparing physics grades with other science grades, Bridgham looked separately at biology and chemistry grades with physics grades being more severe than biology but less severe than chemistry. As Dietrich did not use that breakdown, comparison is not possible. While Bridgham's findings are "mixed" because of the inclusion of other variables (student sex, parental occupation, and availability of second level courses) he did conclude that "ease of grading is related to science enrollments in general." This is in direct opposition to the outcomes of Dietrich's study. Clearly, we are not in a position to validly generalize about the relationship between severity (or ease) of grading in enrollment in physics classes. It is fair to conclude that this phenomenon is not a simple one, and is in need of further research which begins to represent enrollment realistically (i.e., in a longitudinal fashion) and to tap other variables that potentially relate to student enrollment in high school physics classes. In addition to the Dietrich and Bridgham (2) studies, reviews by Rowe (4) and Bates (1) will be used to formulate these suggestions for further research.

From Bridgham's study, it is clear that females are penalized more severely than males, when physics grades are compared to their non-science grades. The same finding was obtained with biology and chemistry grades. Bridgham also found stronger associations between ease of grading and enrollment in the next science course for female than for male students. Bridgham speculated that with males the perceived relevance of science courses to career development may override the ease of grading considerations. As Bridgham's data were from the 1968 graduating class, an inspection of this relationship with more contemporary data is highly desirable. Suffice it to say, further research on enrollment in physics should be inspected with the possiblilty of sex bias clearly delineated.

While Dietrich discriminated among those who dropped out of physics and those who did not enroll in physics, students from both groups are missing the contribution that physics instruction might make to their personal lives. It appears necessary to study most educational problems longitudinally, certainly those related to enrollment

characteristics. One might search for differences in demographic or academic variables with respect to those who: (1) choose not to enroll in physics, (2) enroll in physics and drop out, and (3) enroll in physics and complete the course.

Several limitations to physics enrollment appear to exist, prime among them: (1) the lock step sequence of biology-chemistry-physics and (2) the mathematical skills prerequisite to physics. While the 100 year old sequence of the secondary science courses may need changing, such will occur very slowly. More realistically, a wider spectrum of physics courses, for more than just the college-bound science major, might alleviate this enrollment problem. As such curriculum reconstruction is considered, it is imperative to consider making the new courses less abstract and more concrete and less dependent on lecturing solely and more dependent on other techniques of instruction/learning. Researchers could explore the relation between enrollment and physics course variability, in situations where such exists. As enrollment in physics is usually limited to those who have completed chemistry, another index of physics enrollment might be as a ratio of those "eligible" to take physics (i.e., having completed the "prerequisite"). Bridgham used this index in his study. Similarly, the relationship between physics enrollment and enrollment and/or achievement in mathematics courses needs to be explored.

It is not clear how this issue of "prerequisites" for physics is operationalized, informally or formally, in schools. In some cases, it appears that teachers or guidance counselors schedule students into or out of physics on one pretext or another. Other students enroll in physics because they believe it to be required for college admission. While the mechanisms may vary, it appears that further research should pursue this issue, namely, "Who decides if a student should enroll in physics or not and why?"

In addition to the grading practices of physics teachers, many other characteristics of these key individuals may relate to student enrollment in physics classes. While personality variables are admittedly difficult to assess, the warmth and enthusiasm of a physics teacher may directly or indirectly (through physics students) influence enrollment trends. In an area with sex bias apparently existing, it might be valid to explore the effect of teacher's sex on physics enrollment (recognizing that there are relatively few female physics teachers). While the academic background or experience of a physics teacher might influence the depth and breadth of examples and instruction possible, a more fruitful direction to pursue might be an analysis of the actual classroom behavior with respect to laboratory activities, audio-visual aids, solution of problems, individualization of instruction, and method of evaluation. While obtaining external observers to collect these data becomes extremely laborious and expensive, it appears that students can validly report on the kinds of behaviors.

A last cluster of variables potentially related to physics enrollment is composed of measures of the school and community, such as size of school, socioeconomic status of the community, percent of students entering college, etc. While it is likely more influence will be from

other factors more closely associated with the physics class, the teacher and the students, in most educational research, it is wise to consider as wide a sampling of variables as practical. It is hoped that further research will be conducted to illuminate this phenomenon more clearly.

REFERENCES

1. Bates, Gary C. "Move on the Problem of Physics Enrollments." The Science Teacher, 42:29-30, 1975.

2. Bridgham, Robert G. "Ease of Grading and Enrollment in Secondary School Science, II, A Test of the Model." Journal of Research in Science Teaching, 9:331-342, 1972.

3. Bri!gham, Robert G., and Wayne Welch. "Physics Enrollments and Grading Practice." Journal of Research in Science Teaching, 6:44-46, 1969.

4. Rowe, M.-y Budd. "What's Wrong with Physics - or Physics Teachers?" The Science Teacher, 42:21-22, 1975.

Lawrenz, Frances. "Science Teachers' Perceptions of Their Teaching Skills and Their School Conditions." Science Education, 58(4): 489-496, 1974.

Descriptors--Attitudes, Educational Facilities, *Educational Research, Science Education, *Secondary School Science, *Self Evaluation, *Teacher Attitudes, Teacher Education, *Teaching Techniques

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Judy Egelston-Dodd, National Technical Institute for the Deaf, Rochester, New York.

Purpose

The author surveyed junior and senior high school teachers to identify areas of teaching skills in which the teachers felt a need for improvement from a previously designed list of competencies. Teachers were also asked to rate several aspects of their working conditions.

The actual problem investigated involved a comparison of the responses of the junior high school teachers with those of the senior high school teacher.

Rationale

Teacher self ratings of their skills provides critical input for the design and development of in-service programs to upgrade the competencies of professionals in the field. Administrators and developers of undergraduate teacher education programs also need to identify their strong and weak areas of preparation as perceived by their graduates.

Administrators also need to know teachers' dissatisfactions regarding school facilities and other working conditions. It is assumed that teachers will develop these opinions independently and then present their recommendations (anonymously?) to the administration.

Research Design and Procedure

Although no research design was explicitly proposed, teachers' responses were categorized on the basis of junior vs. senior high school teaching assignments, and group means were analyzed (again no method of analysis was specified) for significant differences for each item contained in both parts of the survey.

The instrument contained items related to demographic data and opinion items for the two topics of concern: teaching skills and school working conditions. The twenty teaching skills items covered: 1) effectiveness in using a variety of classroom presentation techniques, 2) knowledge and ability in subject area, 3) ability to change and organize curriculum, and 4) effectiveness in evaluation tasks.

The clusters of 25 items pertaining to existing school working conditions included: 1) course constraints, i.e., equipment, facilities, and quality and quantity of materials used for student learning, 2) time constraints, 3) space constraints, and 4) personnel constraints, i.e., availability of consultants, secretaries, and assistants.

The science teachers rated themselves on each item of each cluster according to the following Likert scale: 5 = excellent, 4 = good, 3 = satisfactory, 2 = some improvement needed, and 1 = much improvement needed.

Interpretations

The mean ratings for teaching competencies reported in the study were discounted as being somewhat inflated due to the difficulty of self evaluation. The high confidence in ability may also be related to the relatively long experience in the field, to the teachers' holding of degrees within the fields of science, and to the high incidence of participation in NSF Institutes.

The teachers' lack of personal experience with a variety of new teaching techniques was cited as contributory to the low rating for the competency item "Knowledge of Curricular Techniques." Undergraduate programs are advised to provide more background in different techniques.

The highly motivating aspects of audio visual materials, and the need to use effective motivation with junior high level students were cited as explanation for the higher ratings given the item "Effectiveness in audio visual presentation" by the junior high teachers. No explanation was offered for the superior ratings of other items by the senior high group.

The relationship of lower ratings for items "Your teaching load" and "Student Classroom Behavior" was stressed. School administrators were advised to check dissatisfaction within their schools particularly to see if improvements in these areas would improve the quality of instruction.

ABSTRACTOR'S ANALYSIS

Professor Lawrenz has reported on an interesting mixture of issues. Self appraisal of teaching competence and the logical follow-up of diagnostic analysis and prescriptively assigned in-service training for professional development is a neglected feature of the evaluation process in the public schools. The inflated results indicate a need for training in just that process of self evaluation [see Egelston and Egelston (1)].

Working conditions and the motivators and dissatisfiers which interact to form the basis of the perceptions have been reported by Herzberg (2). The specific conditions essential for effective instruction in science have never been validated by researchers to my knowledge.

The teacher variable will always obliterate differences in teaching load, provisions for space, time, equipment, etc. A correlational analysis of the competencies as they were self-rated and the working conditions' ratings might possibly yield some interesting relationships.

The report has some limitations which should be noted. The following comments reflect my own curiosity over omissions and discrepancies rather than a true critical analysis.

The problem statement suggests that the type of learning situations which teachers prefer were investigated yet no results for this item were reported. Since Lawrenz offers specific advice to undergraduate institutions of teacher education and even to NSF Institutes which provide in-service training to teachers, regarding the specific areas reported by teachers as deficiencies in their professional competence, it is likely that these additional data would be useful to readers who may be planning to develop and/or revise such programs.

The problem statement neglected to include a description of the investigation into the teachers' attitudes toward science, nor was the instrumentation reported for this variable. Such incidental findings may have been viewed as part of the demographic data but a correlational analysis with the competency items might yield interesting relationships.

It is likely the author misspoke when she described the sampling process. A stratified random sample of schools within the three regions rather than teachers must have been drawn since teachers were "randomly" selected by their school principal. In reality this selection process may have been less than random as evidenced by the long experience and high academic quality of the teachers in the sample. Perhaps principals selected their "best" teachers in science in the hope that their school would look good in the study. Since no verification of the randomness of the selection process was reported, I am suspicious.

The research design and method of analysis were not specified in the report. The statistics reported were straightforward enough to infer their origin, but the omission of this information detracts from the study.

Table 1 contains an error which is easily identified upon examination of Table 2. The asterisk identifying those items which reached significance in the analysis was omitted from "Effectiveness in audio visual presentation."

In the discussion section Lawrenz refers to the "suggested" relationship between teaching load and teacher opinion of student behavior without supplying post-hoc correlational analyses to support her suspicion. Such action would seem appropriate in light of the four other highly significant differences in the item mean ratings. Perhaps student laboratory facilities which were rated as less than satisfactory by junior high teachers is more related to student behavior than class load is.

It was somewhat surprising to see how unfamiliar teachers reportedly were with innovative curriculum techniques, especially since their strength in other areas would lead the reader to believe that they were in touch with educational journals and/or professional associations. The rank ordering for curriculum techniques and current curriculum matter and subject matter placed them at extreme ends of the ratings for both groups of teachers. I am left wondering if defining these items on curriculum might have helped the teachers respond more positively than merely "satisfactory."

Lawrenz ends with the recommendation that school administrators should check out the validity of the dissatisfaction with school conditions reported, but fails to include standards or references for such standards to make such judgments. The importance is not so much that conditions were in reality that bad, but that the teachers sampled thought they were. Dealing with teachers' feelings of dissatisfaction by means of a survey of "institutional health" would establish a basis for an administrator's plans for improvement. Whether these improvements result in higher quality of instruction or not could be answered by an appropriate evaluation design.

In summery, Lawrenz's findings contribute to the conviction that our science teachers are better trained than ever, that senior high teachers have it "better" than junior high teachers and that school conditions are perceived as in need of improvement by all teachers.

REFERENCES

1. Egelston, Richard L. and Judy C. Egelston. "Self Evaluation and Performance on Classroom Tests," Science Education, 57(4): 525-531, 1973.

2. Herzberg, Frederick. Work and the Nature of Man. Cleveland: World Publishing Company. 1966.

Mayer, Victor J., John F. Disinger, and Arthur L. White. "Evaluation of an Inservice Program for Earth Science Teachers." Science Education, 59(2):145-153, 1975.

Descriptors--Discovery Learning, *Earth Science, *Evaluation, Higher Education, *Inservice Teacher Education, Institutes (Training Programs), Science Education, *Teacher Behavior, *Teacher Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Herbert A. Smith, Colorado State University.

Purpose

The study extended over a three year period and was designed to determine the impact of an inservice program on teacher classroom behavior and effectiveness. The null hypotheses were not explicitly stated, but can be inferred to be:

1. There is no difference in participating teachers' knowledge of earth science facts and concepts as a result of participation in a four-week workshop designed to improve knowledge in this area as measured by pre- and post-test scores on a staff-developed instrument.

2. There is no difference in the learning of the students of participant teachers of factual information and science principles as measured by pre- and post-test scores using TOSK Tests I and II at the beginning and end of the school year.

3. There is no difference in the learning of the students of participant teachers in understanding scientific concepts and processes as measured by the "Concept Process Test" (CPT) scores at the beginning and end of the school year.

A series of hypotheses were tested which related to the changes in perceptions of the participants' students toward classroom activities as measured by the "Science Classroom Activity Checklist: Student Perceptions" (SCACL:SP) as a pre-test, post-test and follow-up test. The plan was designed to measure the impact of the program at the end of either one year or two years of participation and also the retention of program effectiveness on the classroom behavior of teachers as measured one year after program termination. A sample null hypothesis for this series could be inferred to be:

There is no difference in the perceptions of students toward classroom activities methods and techniques used by participating teachers in the classroom as measured by pre- and post-test measures using the "Science Classroom Activity Checklist: Student Perceptions" at the beginning and end of the school year.

Rationale

Educators are constantly faced with the problem of measuring the effectiveness of certain educational procedures. Large sums of money were expended on many aspects of science and mathematics education by the National Science Foundation in the 1960's and 1970's. The principle of accountability requires that a satisfactory assessment of the effectiveness of these extensive programs be made. The present study represents one small attempt to make a contribution to this assessment. A comprehensive assessment and synthesis of the full-scale impact of the National Science Foundation's commitment of large sums to the upgrading of science teacher qualifications and science curricula remain to be done. Blosser is cited in the present study and sufficiently documents the limitations of existing studies. The study by Mayer, Disinger, and White attempts to avoid some of the limitations in previous studies by looking more directly at teacher classroom behavior and student performance and perceptions.

Research Design and Procedure

The study was initiated with 32 junior high school teachers as program participants and included the students in their classes. The study covered two years, with measures of student perceptions of teaching behaviors continued through a third year. The design was traditional and used a series of pre- and post-test measures both on teacher-participants and on the students of participant teachers. Data analyzed were obtained from the pre- and post-test testing procedures and used the "t" test for matched pairs to determine the significance of obtained differences.

Findings

The findings may be summarized as follows:

1. Teachers made significant increases in knowledge of earth science facts and concepts as measured by a locally developed measuring instrument.

2. Students in participating teachers' classes made significant increases in knowledge of scientific facts during the 1969-1970 school year as measured by the TOSK test. Increases in understanding of science principles were not statistically significant as measured by TOSK.

3. Students in participating teachers' classes exhibited significant positive increases in their perceptions of teacher behaviors more consistent with current philosophies of effective science teaching as measured by the "Science Classroom Activity Checklist: Student Perspective." The gains tended to persist during the year following termination of the inservice program.

Interpretations

The investigators believed the evidence indicated that the purposes of inservice program had been achieved. No direct evidence was obtained relating to an objective related to assisting teachers to develop techniques of self-assessment of their classroom behaviors.

ABSTRACTOR'S ANALYSIS

The intent of the study is clearly laudable and it seems to have been carried through consistently and fairly. Nevertheless, there are serious deficiencies which seem apparent to this reviewer.

1. Conceptually, the study is weak and the results obtained could have been predicted a priori with little risk. Certainly one would anticipate that "intensive all-day sessions" conducted on geological content for four weeks would result in teachers learning facts, principles and concepts of geology. Furthermore, from the analyses we cannot separate teacher effects, curriculum effects (traditional versus ESCP) or grade level effects. The statement that "the use of a control group design was not practical" cannot be accepted without substantial reservation. From the study one does not feel assured that teachers who did not participate in the workshop experience would have obtained any less impressive results. In fact, it is somewhat disappointing and a clearly unimpressive finding that students did not show a significant growth in understanding scientific principles in geology after a year of study of earth science. It may show teacher emphasis on, and concerns for, the lower and more superficial levels of cognitive learning as exemplified in Bloom's taxonomy by levels 1, 2 and 3.

2. Apart from the study itself, there is at least the possibility of a conflict of interest in this study. It is noted that the National Science Foundation funded both the inservice program and development of the ESCP material. The Director of the Inservice Project was apparently also the principal investigator in the evaluation study, at least no disclaimer is made. Although no negative allegation is either made or implied, it is suggested that the best interests of evaluation are served when the evaluator has no possible vested professional or financial interest in the outcome of the study.

3. As a technical note on the analysis, it appears that, since it was students and not teachers who were paired, that the "N" in tables III, IV and V should be the student and not the teacher "N". Furthermore, the delay in publication following completion of the study seems excessive.

Perkes, Victor A. "Relationships Between a Teacher's Background and Sensed Adequacy to Teach Elementary Science." Journal of Research in Science Teaching, 12(1):85-88, 1975.

Descriptors--Educational Research, *Elementary School Science, Higher Education, *Instruction, Personality Assessment, Preservice Education, Science Education, Teacher Education, *Teacher Attitudes, *Teacher Background

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Eugene L. Chiappetta, University of Houston.

Purpose

The purpose of the study was to explore the relationship between prospective elementary teachers' backgrounds and their confidence to teach science. The results were expected to support the contention that individuals who took fewer collegiate science courses and who possessed a poor grasp of methods and procedures of science would have a low preference for teaching science.

Rationale

Unrewarding experiences in science courses probably produce poor attitudes toward science and prevent students taking elective science courses. This vicious cycle is the cause of teachers' lack of desire for teaching science in the elementary school. Teachers, like other people, will do what they are successful at and will avoid activities in which they perceive themselves to be inadequate. The apparent lack of science teaching which exists in the elementary school is a result of teachers' attitudes toward science which were shaped during their science course experiences at the collegiate, secondary, and possibly elementary school levels. This notion is supported by the work of Soy (3), Blackwood (1), and Maben (2).

Research Design and Procedure

The sample for the study consisted of 52 prospective elementary teachers enrolled in a teacher education program at the University of California at Davis. Each subject was administered the following three instruments:

(1) A questionnaire soliciting selected biographical information, e.g., the number of high school courses completed, the number of science courses (nationally developed programs) completed, the number of college science courses taken, and a self rating on a sense of adequacy to teach science.

(2) The Omnibus Personality Inventory was used to assess ego-functioning, e.g., social emotional maturity, social concern, success; and to assess "intellectual activity," e.g., working with ideas and abstractions, theoretical orientation, and esthetic interests.

(3) The Methods and Procedures of Science: An Examination, assessed subjects' grasp of methodological dimensions of science, e.g., hypothesis, observations, variables and experimental design.

Correlational analyses were computed for the selected variables using the Bio-Med program. The analyses were for the purpose of answering the following questions: Do individuals with affinity to engage in theoretical problem-solving tasks study more science and have a higher desire to teach science? Does the understanding of the process components of science, an important dimension of contemporary science programs, relate significantly with background variables such as flexibility or the tendency to engage in intellectual activities?

Findings

There was a significant relationship between a prospective elementary teacher's academic experiences in science courses and his tendency to elect additional college level science course work, and the preference and adequacy for teaching science in the elementary school. There was a significant correlation between a prospective teacher's sense of adequacy for science activities and the preference to teach elementary science. In addition, there was a significant relationship between individuals' grasp of the methodological dimension of science and the number of nationally developed high school science programs that were studied.

Interpretations

The study adds more credibility to the notion that an individual's perception of adequacy for a task will shape his attitude for engaging in the task and in related endeavors. Elementary school teachers who have had successful experiences in science course work will be inclined to teach more science in their classrooms than those teachers who struggled through their science courses. Unfortunately, the latter probably comprises the majority of cases. Thus, we can expect very little science teaching to occur in elementary school classrooms.

A high relationship was reported between scores on the Methods and Procedures of Science: An Examination instrument and the number of nationally developed science courses that were taken in high school. This suggests that the newer high school science programs have imparted to this sample a grasp for the methodological dimensions of science. This should not come as a surprise because a major goal of the national curriculum projects is to enhance inquiry and process skill development.

A discouraging result occurred in that the prospective teachers did not appear to realize the significance between being able to grasp the methods of science and effective science teaching. The individuals were inclined to associate science content with effective science teaching. This is problematic when we consider the process and inquiry emphasis that has been set forth by the developers of the most innovative elementary science programs produced over the past fifteen years.

There was not a significant relationship between personality attributes and attitude toward teaching science. This is rather surprising when we consider the associations that have been reported between personality dimensions and disposition toward various tasks that have been reported in the research literature. This questions the suitability of the instruments used to measure these associations.

In summary, prospective elementary teachers' perceptions of their adequacy for teaching elementary school science appears to be related to their success in previous science course experiences. Previous science course experiences which imparted a factual grasp of science appear to be associated with a sense of adequacy for teaching science. This is unfortunate because a methodological grasp of science might better be the desired association with adequacy for teaching science in the elementary school.

ABSTRACTOR'S ANALYSIS

The study adds to the fund of research which supports the notion that teachers will be inclined toward those activities for which they feel a sense of adequacy. The reported finding, that experiences which promote a grasp of the factual content of science might influence teachers' understanding of what adequate science instruction is, should be pursued in subsequent research. This might be instrumental in formulating a hypothesis which suggests that an understanding of what adequate elementary science instruction should be like involves a two-stage development. The first stage involved successful experiences with the mastery of science content, while the second stage involves successful experiences with and mastery of science process skills.

Most prospective and practicing elementary school teachers never achieve the first stage, not to mention the second stage. Possibly, only after equilibrating at the second stage will elementary teachers carry out the goals of the nationally developed elementary science programs and continue this pursuit over an extended period of time.

The findings in this study were extremely difficult to understand. There was no results section, per se, and no tables. Without the results section one could not be sure of the significance of the reported correlations. The strength of relationsips between variables cannot be ascertained. This study, presented in its research report form, is highly fragmented. Hence, the reader cannot determine what was found nor accurately understand the interpretations of the findings by the author. Research reports will be more helpful to the readers if they contain a results section with tables that show relationships between the key variables.

The present study might have employed a stronger design than what appeared to be employed. The attempt to find correlations among so many (21) variables was sure to meet with some type of success. The study would be more valid if fewer relationships were analyzed using a regression analysis or another multivariate statistic.

REFERENCES

1. Blackwood, Paul E. "Science in the Elementary School." School Life, 47:13-28, November, 1964.

2. Maben, Jerrold W. "Science Teaching in Elementary Schools." The Science Teacher, 40(7):40-41, October, 1973.

3. Soy, Eloise. "Attitudes of Prospective Elementary Teachers Toward Science as a Field of Study." School Science and Mathematics, 67:507-517, 1967.

Sayer, Irwin, James R. Campbell, and Cyrus W. Barnes. "The Effect of College Instructors' Interaction Ratios on Cognitive Development." Science Education, 56(4):529-537, 1972.

Descriptors--*Academic Achievement, *Biology, College Science, Educational Research, *Instruction, Science Education, *Student Teacher Relationship, Teacher Characteristics

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Dorothy Gabel, Indiana University.

Purpose

The purpose of this study was to examine the relationship between the styles of instruction of teachers of introductory college biology utilizing Flanders' Interaction Analysis System and student achievement on both high and low cognitive levels. (High cognitive level included Bloom's comprehension, application and analysis categories; low cognitive level, Bloom's knowledge level.)

Rationale

Within the past 40 years many research studies have focused on the effectiveness of specific teaching methodology for college instruction. No clear patterns have emerged that favor any specific methods. At the time this research was conducted, studies that examined different teaching strategies failed to quantify the method used. For example, "student centered" in one study might mean letting the students talk 20 percent of the time, whereas in another study it may mean 100 percent of the time. This study attempted to counteract this weakness of previous studies by quantifying the teaching approach used (lecture vs. student-centered) utilizing the methodology developed by Flanders.

Research Design and Procedure

The sample consisted of 408 students enrolled in an introductory biology course at Bronx Community College in New York City in 1967. Half of these students were enrolled in classes taught by instructor-centered teachers, and half in classes taught by student-centered teachers. The teacher's classification was made by analyzing the teaching behavior of all instructors of the introductory biology course according to Flanders' Ten Category Interaction Analysis System and selecting those with ratings at the extremes of the "direct" and "indirect" scales.

Students were administered a pretest (not described in article), were taught for one semester by a "direct" or "indirect" instructor, and took the posttest at the end of the semester. Instructors were videotaped for five one-hour sessions during the semester. These tapes were analyzed by the major investigator according to Flanders' system. (Care was taken to determine consistency in ratings using the instrument and inter-observer reliability).

The posttest consisted of 104 items equally divided between low and high level questions. Validity of the instrument was determined by two panels of experts who evaluated items on the quality, cognitive level, and content covered. The reliability of the test using the Spearman-Brown formula was 0.80.

Data were analyzed using a series of one-way analyses of covariance. Pretest and IQ inequalities were used in adjusting the final posttest scores.

Findings

Analysis of the teachers' classroom behavior according to the categories established by Flanders indicated that there was a significant difference at the 0.01 level between direct and indirect instructors' methods of teaching. Comparisons were made between I/D, i/d, I/I+D, i/i+d, and $I/D_{8,9}$. The disparity between the direct and indirect groups was greatest for the I/D ratio indicating great differences between the groups in questioning and lecturing behavior. (The indirect group used 18 percent of its time in questioning, while the direct group used only 2 percent.)

Although there was a difference in the treatment for the two groups of students as shown above, this difference in treatment did not result in significant differences in the way students performed on the biology posttest. Scores for low level achievement, high level achievement, and total achievement proved to be nonsignificant. Because means were not listed in the article, no judgment can be made concerning trends in the scores.

Interpretations

The authors of this research report concluded that although two methodologies were clearly delineated and found to be significantly different, neither method was more effective in terms of achievement. As a final segment of this study, grand matrices produced in this study were compared with those reported in the literature for other groups of students. A comparison of scores among junior, senior high, and college science instructors indicated that each level exhibits distinct interactive behavior. Perhaps this is the reason why indirectness has been found to be effective at the junior high level but has not been found to be as clearly related at higher grade levels. Other instruments may be needed to uncover variables related to teacher effectiveness at the college level.

ABSTRACTOR'S ANALYSIS

One of the major weaknesses of studies of teaching methodology is failure to monitor classroom procedures to assure that the teaching strategy being examined is properly executed. An attempt was made in this study to prevent this flaw by videotaping a sample of the instruction. In addition, the researchers tried to carefully define "teacher-centered" and "student-centered" instructors by reporting their I/D ratios. This certainly aids other researchers in replicating the study.

Another strength of this study that must be mentioned is the careful and methodical establishment of the validity and reliability of the

instruments used. The researchers made every effort to establish the validity of the analysis of teacher behaviors from the videotapes by determining both the consistency of the rater on the same tape and the correspondence of this rating with that of others. The achievement test appeared to be carefully constructed using a panel of instructors to judge suitability of the content and a panel of experts to judge the cognitive level and quality of the test items.

Several aspects of this study are not clear from the written report.

1. In the description of the sample, it states that there were 204 students in each group. No mention is made if these students were randomly assigned to the treatments or if they were equally distributed among the six instructors. A more complete description of sample selection is necessary in order to determine whether the investigation is a true experimental study.

2. The researchers utilized a pre- posttest design with an analysis of covariance to adjust for any inequalities in the sample. No description is given of the pretest. Is it the same as the posttest? If not, how does it correlate with the posttest?

3 The article lacks sufficient tables for the reader to make judgments about the outcome of the experiment. Two tables are included in the article but these do not list sample size, means, or standard deviations. It is not possible from the report to tell whether the teacher-centered or student-centered instruction effected higher mean scores. The reader only knows that the differences were significant at the 0.2 and 0.3 level.

Because of the inadequate description of the sample, it is difficult to determine whether the analysis of the data was performed properly. It would appear if six instructirs were used in the study that students were in small classroom groups of approximately 70 or less students. (The size of the group may be an important variable that is not considered in this experiment.) The proper unit of analysis should be the classroom if this is the case, rather than the individual student.

A factor that the authors did not discuss that may have accounted for no significant differences between groups was the low number of hours that the instruction was monitored (five one-hour sessions). It may have been possible that the instructors behaved in a certain manner while the class was being taped that was not their usual style, and that there really was no great difference among instructors. The study could have been strengthened by videotaping a large number of sessions and randomly selecting five hours from those taped. The description of the taping is insufficient in the article to know if this was the case.

Finally, the authors of this study set out to determine if there were cognitive effects in teaching students using a teacher-centered versus a student-centered approach. They found no differences in biology achievement. It may have been an interesting addition to this study to check students' attitudes toward biology in using the two approaches. This finding may be of equal importance to cognitive growth.

STUDENT CHARACTERISTICS

AND BEHAVIORS

Horn, Jerry G. "Risk-Taking in Explanation of Biological Events." Journal of Research in Science Teaching, 10(4):341-346, 1973.

Descriptors--Biology, *Cognitive Processes, Educational Research, *Group Relations, *Individual Characteristics, *Risk, Science Education, Secondary School Science, *Student Behavior

Expanded Abstract and Analysis Prepared Especially for I.S.E. by David R. Stronck, Washington State University.

Purpose

The purpose of this investigation was to determine correlations between six different circumstances and the risk-taking behaviors of biology students as they attempted to explain observed biological events. The six different circumstances provide the following six questions considered by the study: (1) Do individuals differ in their risk-taking behaviors in relation to their degree of achievment motivation? (2) Do individuals acting as a group express more risk-taking explanations than members of the group acting separately? (3) Is there a shift toward more risk-taking explanations by individuals after participating in a small group discussion? (4) Do discussion groups formed on the basis of similar or dissimilar degree of achievement motivation exhibit greater risk-taking behaviors than the individuals of the group acting separately? (5) Do individuals and groups exhibit greater risk-taking behaviors in explanation of situations that are open to argument as compared to those that are not (explanation unknown vs. explanation known)? (6) Is there a correlation between risk-taking and the individual's IQ or sex?

In this study "risk-taking" is defined by the presence of the following behaviors: (a) exhibition of extremity and confidence of judgment in situations where greater extremity affords the possibility of greater magnitude of error; (b) free participation in discussion by expressing his or her opinions, feelings and/or criticisms regardless of the presence of peers and/or authority; (c) willingness to try new approaches in explanation of observed events.

"Achievement motivation" is defined as the striving of an individual to achieve academic success, as measured by opinions and self-reported activities recorded in the Achievement Motivation Test. This test was a modification of the one reported by Russell (4). The Hoyt reliability of the modified form was 0.62 and the test-retest reliability over a two-month period was 0.83. Students were placed in the category of "high achievement motivated" when they were in the top one-third of the total sample or had scores equal to or greater than the raw score closest to the 66 2/3 percentile on the Achievement Motivation Test. The category of "low achievement motivated" was for students in the bottom one-third of the total sample or scoring equal to or less than the raw score closest to the 33 1/3 percentile on the Achievement Motivation Test.

The "biological event" observed in this study was one of the two sequences in the film, Elephant Seals, prepared by the Biological Sciences

Curriculum Study. These sequences show two male seals in physical conflict and the reaction of one seal to the recorded sounds of another seal.

Rationale

A cooperative project of the Mid-Continent Regional Educational Laboratory (McREL) and the Biological Science Curriculum Study (BSCS) identified behaviors related to inquiry. One of these behaviors is risk-taking. The literature of psychological research defines risk-taking as a person's willingness to gamble for extrinsic rewards in games of chance. Nevertheless, the project by McREL and BSCS defines risk-taking in terms of a person's willingness to express opinions in class discussions regardless of the presence of authority and possible criticism. This latter type of risk-taking is needed to allow class discussions of biological events to become inquiry lessons. This study seeks to identify the circumstances which tend to promote increased risk-taking and therefore more fruitful inquiry lessons.

Research Design and Procedure

This study does not have an experimental design. The researcher did not attempt to change the behaviors of students but rather to identify correlations between the behavior of risk-taking and various circumstances. Two different measures of risk-taking were used in this study: the Extremity-Confidence of Hypothesis Test and the Risk-Taking Verbal Observation Scale. The reliability of the Extremity-Confidence of Hypothesis Test was found to be 0.60 using the Spearman-Brown Prophecy Formula. The Risk-Taking Verbal Observation Scale had its reliability estimated by the consistency of categorization as suggested by Cronbach (1). On a sample of 42 statements, all observers agreed on 84 percent of the items; two-thirds agreed on the remaining 16 percent of the items.

The subjects for this study were members of six high school biology classes that were randomly selected from a large city school district. All of the students were using the textbook BSCS, Green Version, and were in "regular" biology classes, as opposed to accelerated or remedial type courses. The mean age of the total sample (160 males and 156 females) was 16.2 years. Their mean IQ was 109.39 with a standard deviation of 11.90.

The subjects were shown the film, Elephant Seals. Two scenes in the film were used by the researcher as the primary focus for obtaining the students' responses. By random procedures, the subjects were assigned by classes to two knowledge conditions of the status of the information. One group was told that information was known about the behaviors or events they observed. The other group was told that the knowledge condition was unknown and therefore their opinions or interpretations would not be contradicted.

The Extremity-Confidence of Hypothesis Test was administered to each student. On the basis of their scores on the Achievement Motivation

Test, the students were assigned to one of three different types of discussion groups: like-high, like-low, or unlike. One judge per group using the Risk-Taking Verbal Observation Scale observed the students during a 10-minute discussion of the two scenes. At the end of the discussion period, the group completed a second copy of the Extremity-Confidence of Hypothesis Test that gave the group's decision. Then the groups were disbanded. A third administration of the Extremity-Confidence of Hypothesis Test again requested the individual's responses. This test was used to measure individuals twice and groups once.

The dependent variable of risk-taking (as measured by the Extremity-Confidence of Hypothesis Test and the Risk-Taking Verbal Observation Scale) was analyzed within factorial designs: (1) 2 X 2, and (2) 2 X 3. The 2 X 2 factorial design consisted of two levels of achievement motivation and two levels of status of information and a three-level variable according to the grouping for discussion by (1) like-high, (2) like-low, and (3) unlike.

To increase the possibility of rejecting the null hypotheses, the 0.10 level of significance was chosen. Various statistical analyses were performed to recognize significant differences between means and correlation coefficients.

Findings

In this study the Extremity-Confidence of Hypothesis Test failed to measure any significant differences. Nevertheless, the Risk-Taking Verbal Observation Scale uncovered some significant differences. For example, the risk-taking involved in verbal discourse by individuals was affected by differential knowledge of the information. It was concluded that when the information or explanation concerning a biological event is said to be unknown or unexplained, individuals tend to exhibit greater risk in regard to verbal discourse as opposed to individuals who are told that the event has been explained.

Another conclusion is that the grouping arrangement for discussion does affect risk-taking by the group in verbal discourse. It was found that groups like-high and unlike both are significantly greater in risk-taking than like-low. Like-high did not differ from unlike. Because individuals did not differ under the same conditions, it is evident that the group's composition does influence risk-taking.

Based on the correlational analyses, it is concluded that risk-taking is not a function of the sex of the individual. Risk-taking in verbal discourse is only slightly, but significantly, positively correlated with the IQ of individuals.

Interpretations

There was no significant difference found between individuals and groups in risk-taking as measured by the Extremity-Confidence of Hypothesis Test. This test under the conditions of this study lacked the power of

detecting differences which probably exist. Research studies, expecially by Wallach, Kogan, and Bem (5, 6) have consistently reported groups exhibiting greater risk than individuals. They explained these outcomes as the diffusion of responsibility among the group members.

The Risk-Taking Verbal Observation Scale detected significant differences by which risk-taking involved in verbal discourse increased under the following circumstances: (1) when the information or explanation concerning a biological event is said to be unknown or unexplained, (2) when the students with high scores on the Achievement Motivation Test are grouped together or mixed with other students, and (3) when the students have higher IQ scores. These conclusions are not sufficient to establish the nature of risk-taking, that is, whether it is a general trait or a multi-dimensional trait.

ABSTRACTOR'S ANALYSIS

The researcher clearly recognized the study of risk-taking in the literature of psychology. He attempted to pioneer science education research in the consideration of risk-taking. His inspiration for formulating the study was based on the suggestion derived from Inquiry Objectives in the Teaching of Biology, Position Paper of the Mid-Continent Regional Educational Laboratory and Biological Sciences Curriculum Study (3).

Unfortunately the use of the term "risk-taking" by this Position Paper significantly differs from that found in the literature of psychological research. While the psychologists have focused their attention upon the motivations for taking risks, this study provided significant data only in terms of the willingness to express opinions in verbal discourse. Because all of the students involved in this study were approximately 16 years of age, the behavior of these teenagers in verbal discourse must be interpreted in the context of adolescent psychology. Typically these teenagers are very conscious of their social interactions with their peers. This study identified some of the traits of individuals who are more aggressive in expressing their opinions in verbal discourse. The abstractor recognizes that this behavior is helpful in performing an inquiry lesson with a BSCS film. But the abstractor suspects that the original concept of "risk-taking" as described by the literature of psychological research may suggest more profound insights into science education.

Thomas S. Kuhn, in his book The Structure of Scientific Revolutions (2), observed: "Bringing a normal research problem to a conclusion is achieving the anticipated in a new way, and it requires the solution of all sorts of complex instrumental, conceptual, and mathematical puzzles. The man who succeeds proves himself an expert puzzle-solver, and the challenge of the puzzle is an important part of what usually drives him on." Probably the motivation of scientists should be analyzed in terms of their willingness to risk their energies for the possible solution of a scientific puzzle. Many teenagers have generously dedicated their efforts to a science fair project or a research topic for a science talent search. This behavior seems to be associated with "risk-taking" as defined in the literature of psychological research, that is, a

willingness to gamble for extrinsic rewards in games of chance. It seems less associated with "risk-taking" as defined by the McREL and BSCS Position Paper, that is, a willingness to express one's opinions, feelings, or criticisms regardless of the presence of authority and a willingness to participate freely in class discussions.

The abstractor suggests that a very important problem for science education research is the identification of the circumstances which favor risk-taking, that is, the willingness to seek the solution to an involved scientific puzzle. Unfortunately there may be only a weak relationship between this type of risk-taking and the willingness to interpret in verbal discourse the behavior of Elephant Seals after viewing a film for a few minutes. The discussion period of ten minutes may have been too brief to allow some thoughtful students to express their insights. The social dynamics required in verbal discourse may have obscured the more profound question concerning the student's willingness to take the risk of a scientific investigation.

The subjects were requested to interpret the behavior of Elephant Seals. This topic is not typical of those considered in a high school biology class. The behavior of the seals can be related to human psychology and sociology. Certainly the subjects who were Midwestern teenagers were unfamiliar with observing either the behavior of Elephant Seals or similar animals. Perhaps a better topic would have been one of plant physiology, that is, a topic which could relate to students' previous observations but which does not overlap with the human social sciences. The exclusive use of the film, Elephant Seals, raises some important questions on the validity of this study. A more typical biology topic might have generated significantly different responses from the subjects.

The researcher was unable to discover any significant differences by using the Extremity-Confidence of Hypothesis Test. This conclusion could have been anticipated by noting that the reliability of this test is only 0.60. Moreover each subject was required to complete the same test three times within a relatively short period of time. Probably the teenagers in this study tended to repeat their responses by memory when they were asked to complete the test for the second and third time. The abstractor does not anticipate that this test used with the procedures of this study will provide significant differences although the topic and subjects may be changed.

The current state of research on defining "risk-taking" in science education remains embryonic. The researcher recommends "the replication of this study using an instrument for extremity-confidence of judgment and larger sample sizes, especially for group data, that would provide greater power for detecting a difference." The abstractor recommends that future studies should move in a different direction. The emphasis upon verbal discourse should be replaced by an emphasis upon the selection of open-ended laboratory experimentation. This latter emphasis will clearly relate risk-taking to the most important processes of science, e.g., gathering data to support an hypothesis. The scientific topics considered should be of the type which can be investigated in a typical high school laboratory. The appropriate tests for measuring

the students' performances should consider not only the willingness to generate hypotheses but also the formulation of reasonable procedures for gathering data which may be relevant to supporting the hypotheses. Such recommended studies may provide information on the circumstances which encourage adolescents to do science, that is, to take risks of the same type as those selected by professional scientists.

REFERENCES

1. Cronbach, Lee J. Essentials of Psychological Testing, 2nd Ed. New York: Harper and Bros. 1960.

2. Kuhn, Thomas S. The Structure of Scientific Revolutions, 2nd ed., enlarged. Chicago: The University of Chicago Press. 1970.

3. Mid-Continent Regional Educational Laboratory and Biological Sciences Curriculum Study. Inquiry Objectives in the Teaching of Biology, Position Paper, 1 (1), September 1969.

4. Russell, Ivan L. "Motivation for School Achievement: Measurement and Validation." The Journal of Educational Research, 62:263-266, 1969.

5. Wallach, M., Nathan Kogan, and D. J. Bem. "Group Influences on Individual Risk Taking." Journal of Abnormal and Social Psychology, 65:75-86, 1962.

6. Wallach, Michael, Nathan Kogan, and D. J. Bem. "Diffusion of Responsibility and Level of Risk Taking in Groups." Journal of Abnormal and Social Psychology, 68:263-274, 1964.

Koelsche, Charles L. and S. Lloyd Newberry. "A Study of the Relationship between Certain Variables and the Science Interest of Children." Journal of Research in Science Teaching, 8(3):237-241, 1971.

Descriptors--Educational Research, Elementary School Science, *Interests, Race, Science Course Improvement Project, *Student Science Interests, Sex Differences, *Student Characteristics

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Hans O. Anderson, Indiana University-Bloomington.

Purpose

The first purpose of this study was to construct and validate an instrument for determining children's science interests. Second, this instrument was used to compare science interests of selected students in grades four and six from two schools utilizing different methods of science instruction. Four null hypotheses were tested:

1. There is no significant difference in the science interest categories of children enrolled in the "Process Approach" science curriculum and those enrolled in classrooms where the "Process Approach" does not exist.

2. There is no significant difference among the science interest categories of children enrolled at the fourth and sixth grade levels.

3. There exists no significant difference in the science interest categories between boys and girls.

4. There exists no significant difference in the science interest categories between Negro and Caucasian children.

Rationale

The authors noted that pupil interest is recognized as a main factor influencing learning. Curriculum developers and teachers planning classroom science experiences continually strive to develop experiences that will be interesting to students. Therefore, it follows that identification of student interests and of the potential effects of selected variables is necessary; and that valid and reliable instruments are needed for this purpose. The authors extend their rationale by pointing out the importance of determining whether certain selected variables differentially relate to interest.

Research Design and Procedure

The study involved instrument development, validation, and hypotheses testing.

Instrument Development

The What I Like to Do Science Inventory is a forced choice instrument composed of 36 pairs of items which permit responding students to select most interesting statements from a nine-category classification system. Each item is matched against each possible choice, forming 36 pair choices. The authors developed the instrument, and it was administered to a population of 100 fourth and sixth graders to obtain estimates of the instrument's reliability and validity.

Two methods of establishing reliability were employed. First, all of the "A" statements from each of the "A-B" pairs on Form A were tallied and compared with the number of "A's" chosen on Form B. This yielded an overall coefficient of correlation of .89. Second, a rank difference correlation of each child's responses on Form A and Form B was calculated. The "A-B" correlations ranged from a low of .63 to a high of 1.0 with a median of .917 and a mean of .908.

The validity of the instrument was tested by correlating the children's interest as measured by the instrument with the parents' statements of what they thought would be their child's interest. The authors indicate that 39 out of 48 participating sets of parents indicated agreement with their respective children. The authors reported a chi square value significant beyond the .001 level. The authors concluded that they had developed a reliable and valid rating scale.

Hypotheses Testing

In part two of the study, the four hypotheses were tested to determine if the variables of curriculum, sex, grade level, and race were discriminating factors among these interests. The analysis of the data included the Multivariate Analysis of Variance (MANOVA) which provided tests for the overall effects of each design factor through a series of two-way classifications. The design used was a factorial analysis of variance design.

Findings

The investigators reported the following findings. (Part II only.)

1. There were significant differences in interest that could be attributed to curriculum, grade level and sex.

2. There were no significant interest differences that could be attributed to race.

The specific nature of these differences is summarized in this table extracted from the article.

TABLE III

Correlations between Discriminant Functions and Original Variables

Science Areas	Curriculum*	Grade*	Sex*
1. Rocks and Soils	.215	.501**	.443**(i)
2. Earth and Space	.042	.402**	.396**(i)
3. Light	-.449**	.016	-.102
4. Heat	-.540**	-.330	.109
5. Sound	.211	-.100	-.434**(i)
6. Electricity and Magnetism	-.138	-.547**(i)	.274**(i)
7. Living Matter	.511*	.314	-.241**
8. Matter and its Changes	-.288	.000	-.046
9. Air and Water	.430*(i)	-.227	-.521**(i)

*Curriculum: (-): 6th grade predominance, (+): Non-AAAS. Sex: (-) Female Predominance, (+) Male. Entries not marked negative (-) are considered positive (+).
**Indicates that the effect for this variable was significant at the .01 level, based on the univariate ANOVA's. (i) Indicates that the effect for this variable was significant at the .01 level due to interaction with another factor, and that interpretations of main effects should be done with caution, if at all.

Interpretations

Curriculum, grade level, and sex were found to significantly and differentially affect student interest as measured by the investigations instrument--the What I Like to Do Science Inventory. Race did not appear to either significantly or differentially affect student interests as measured. Interest differences among children were measurable, and interests were found to relate to identifiable factors.

ABSTRACTOR'S ANALYSIS

As Ralph Tyler (2) stated so economically, "Interest is the point of departure." An interested child can overcome all sorts of barriers to learning! An uninterested child may not even try. Then on the practical side, teachers are frequently heard saying, "He's not interested in anything," or, "How can I get him interested?" Interest is certainly the point of departure and we really know very little about the real interests of children. Hence, this research was conducted (in 1968) in an arena that was, and continues to be, a significant line of research. At least, it is significant to those curriculum developers willing to bend the curriculum to meet the child.

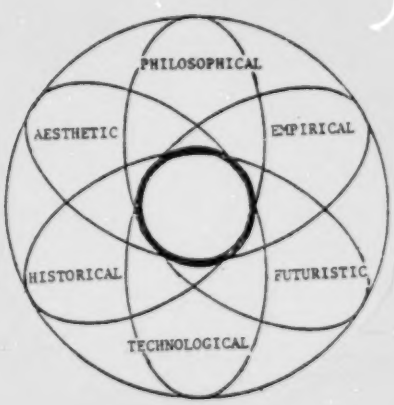
The investigators, like all those involved in test development, were faced with the challenge of attempting to establish instrument validity.

Reliability is a relatively easy, almost concrete concept. Validity can only be hypothesized. In this study the investigators chose a standard approach to establishing validity. They asked parents to indicate the interest of their child and correlated this with the child's test score. Admittedly, parents should know the interests of their children, or at least be able to rank interest preferences given choices. But, is that really the case? My children have convinced me that if I'm not dead wrong about their interests, I'm at least out of date. But certainly, parents are probably better predictors than some outside agents.

It's always easier to correct somebody else's study (frequently because you are not aware of necessary compromises) and it's always advantageous to get in the last word because through listening you can take advantage of and utilize all the good thinking preceding yours. The validity established in this study through reliability studies and correlating parent opinion of their child's interest is impressive. But, would it not be nice if the investigators could have established predictive validity--validity estimates that would permit us to predict that child "X" would be interested in "Y" activity?

In my beginning sentence I quoted Tyler--"Interest is the point of departure," which in this case seems most appropriate. The investigators obviously are of similar ilk! Would it not have been useful to set up a curriculum cafeteria in which they could have observed the child to determine predictive validity? Child"X" is interested in "Y" type activity as measured by the Science Interest Inventory, and, when given the opportunity in a free setting, he pursues a "Y" type activity.

Predictive validity is one point. Another is the factors studied--sex, curriculum, age, and race. The study was conducted in 1967 and it is already ten years later. If I were to replicate this study I would probably not look at any of the variables mentioned (partly because I read this article). The major thesis is student interest and the assumption is that interest generates task orientation and learning. Think for a moment about the curriculum--the stuff we think we ought to teach. What are its interest dimensions? What is it that will interest students? Where and what are the appropriate points of departure? Examine briefly the following model.



The dimensions outlined can be viewed as significant dimensions for the study of science or, for that matter, any discipline. Each lobe represents a significant dimension of study. Each lobe may be a crude representation of the interest orientation of students.

Two significant points to remember that were demonstrated in the study are that curriculum and the sex of students influence student interest. An extension of this research that would examine student interest in respect to other models could be most useful, particularly if the findings would then be used to develop a variety of curricula.

The model suggested above is patterned after the model used by Allport, Vernon, and Lindzey in their 1931 Study of Values (1). It is suggested because the content model used in this study may not provide much guidance to curriculum developers. It is unlikely that many children have inherent content orientations. But children and adults are found to be oriented toward the types of activities inferred by this value-oriented model.

REFERENCES

1. Alport, Vernon, Lindzey, Study of Values, Houghton-Mifflin, 1931.

2. Tyler, Ralph W., Basic Principles of Curriculum and Instruction, University of Chicago Press, 1949.

Peterson, Rita and Lawrence Lowery. "The Use of Motor Activity as an Index of Curiosity in Children." Journal of Research in Science Teaching, 9(3):193-199, 1972.

Descriptors--*Behavior, *Curiosity, Elementary School Students, *Motor Reactions, Racial Differences, Rating Scales, *Research Techniques, Student Characteristics

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Ann C. Howe, Syracuse University.

Purpose

The purpose was to investigate curiosity in children of elementary school age by (a) comparing children by age, sex, and race on a measure of curiosity based on motor activity and (b) comparing scores on the motor activity measure with scores on a verbal measure with teacher ratings.

Rationale

The investigators believe that one goal of science education should be to preserve and nurture curiosity and that a broader understanding of children's curiosity will be useful in attaining that goal.

A major assumption underlying the study was that curiosity in children may be measured by the number and kind of motor responses directed toward objects. In humans, curiosity has been measured primarily through the use of verbal and representational means; in animals and very young children curiosity has been measured, of necessity, by the observation of exploratory behavior or motor response to objects and events. The exploratory behavior of older children has been studied by several investigators, including the authors of this paper who had previously developed a scale for classifying "curiosity behavior" of first grade children. This paper is a report of the use of that scale with another sample of children.

Research Design and Procedure

Sample. Children in four intact classes, one each of kindergarten, second, fourth, and sixth grade, participated in the study. For classes with an enrollment below 30, additional children were chosen, bringing the total number to 120 children. The children were fairly evenly divided between male and female, black and non-black.

Procedure. Children were informed that they would be interviewed but were told that they might decline to participate. (Two children declined.) Each child was told upon arriving for the interview that there would be a short delay until the interviewer finished some paper work. During this time an assortment of objects, including books and a small animal in a cage was available in the room. All motor activity

directed toward the objects and all questions asked during the 10-minute period were recorded. An interview followed.

Measures. Three measures of curiosity were used. (1) Curiosity Index of Motor Activity (CIMA). Motor responses directed toward objects in the 10-minute period described above were categorized according to whether the child approached, manipulated, or reorganized the object(s). (2) Curiosity Index of Verbal Behavior (CIVA). The number of unsolicited questions asked by each child, excluding repetitious and procedural questions, was treated as a raw score. (3) Teacher Rating scale (TR). Each child was ranked by his own teacher according to a set of instructions developed elsewhere.

Analysis of Data. (1) CIMA (motor) scores were subjected to analysis of variance, using age, sex, and racial-ethnic origin as independent variables. (2) A t-test was performed on the CIVA (verbal) score means of high-CIMA scorers as against low-CIMA scorers. (3) The Spearman Rank Correlation Coefficient was used to compare CIMA scores of those kindergarten children who asked questions. (4) The Spearman Rank Correlation Coefficient was used to compare CIMA scores with Teacher Ratings for each grade level. (5) A t-test was performed on CIVA (verbal) score means of children who were high (above the mean) on Teacher Rating as against children who were low (below the mean) on Teacher Rating.

Findings

The authors report the following findings:

1. Black children showed a greater amount of curiosity expressed as motor response than non-black children.

2. Curiosity increased for black males and decreased for non-black males from kindergarten through sixth grade. There was no corresponding difference for females.

3. Curiosity was not related to age or sex except as noted above.

4. There was a negative correlation between curiosity expressed as motor activity and curiosity expressed as verbal activity.

5. No association was found between curiosity expressed as motor activity and the teacher rating of curiosity.

6. Children who were ranked higher in curiosity by teachers asked more questions (i.e., had a higher CIVA score) than those who were ranked low by teachers.

Interpretations

The investigators note that curiosity as expressed by motor response is stable throughout the age range studied and suggest that curiosity should be encouraged through the provision of materials and time for

exploration during the years spent in elementary school. They speculate on possible factors which might explain the difference in behaviors between black and non-black boys and suggest several possibilities for further study.

The negative correlation between question-asking and motor response to objects was discussed. The authors suggest that perhaps teachers take a positive view of curiosity expressed verbally but that they take a negative view of curiosity expressed as motor activity. Thus, children who ask many questions may receive a high ranking on curiosity but children who approach, manipulate and rearrange objects in the classroom are viewed negatively or ignored.

ABSTRACTOR'S ANALYSIS

This article raises several interesting questions and draws attention to a dimension of children's behavior which is often alluded to but has been infrequently studied by educational researchers. The study should serve its readers by stimulating them to redefine curiosity, to recognize it when they see it, and to be sensitive to the various forms it may take. Small children who ask "Why?" at every turn are usually thought to "have lots of curiosity" but the child who sits, quiet and unnoticed, while he explores grandma's handbag is apt to be told that he is naughty. This paper makes us ask whether theachers are reacting to their pupils in the same way. Do teachers assume that all "why?" questions are signs of curiosity while exploratory behavior is misbehavior? Or do they recognize that "why?" questions may be only an easy way to distract the teacher and "why do we have to do this?" is more often a complaint than a show of curiosity?

Although interesting questions are raised, two significant omissions weaken the paper. The first omission is sufficient detail pertaining to the measures employed. None of the measures is described well enough for the reader to make a judgment as to its reasonableness. Previous publications are referenced, but in the case of little-known measures of this kind the reader is entitled to more information. In an essentially exploratory study rigorous design and testing of instruments for validity and reliability are not expected, but one needs to know more than is given here. For example, how were the three different levels of curiosity weighted in arriving at the CIMA score? Was the number of questions asked during the 10-minute waiting period used as the CIVA score or were unsolicited questions asked during the interview? What were teachers asked to consider in ranking children? These and other unanswered questions make it difficult to judge the significance of data gathered with the instruments.

The second omission is, simply, the data. Where are the mean scores on the CIMA of the groups by class, sex, and race? How much did the black males' exploratory behavior increase and how much did the other boys' exploratory behavior decrease? These data are essential to our understanding of the study. When these data are provided, anyone who is interested may verify the results by analyzing the data as the authors did or in other ways; without these data the reader has no

way to judge the validity of the conclusions drawn. The data given in the two tables do not make up for this omission. Graphs showing changes from one grade to the next would have been particularly helpful.

In assessing the significance of the negative correlation between teacher ratings of curiosity and motor behavior, coupled with the positive relationship between teacher ratings of curiosity and verbal (question-asking) behavior, it is well to bear in mind that the teachers' ratings were global judgments and that the CIMA and CIVA scores were measures of specific behavior in a short time period. Nevertheless, these findings raise interesting questions about teachers' perceptions and possible value judgments of children's motor and verbal responses in the classroom. At the least, there appears to be some confusion or lack of consensus as to what constitutes curiosity.

This study should be seen as an exploration into an area which has interesting implications for the theory and practice of science teaching in the elementary school. We are indebted to the authors for exploring a novel area and for raising interesting questions.

Simmons, Jack and William Esler. "Investigating the Attitudes Toward Science Fostered by the Process Approach Program." School Science and Mathematics, 72(7):633-636, 1972.

Descriptors--*Attitudes, *Discovery Processes, *Educational Research, Elementary Grades, Elementary School Science, *Instruction, *Student Attitudes, Student Characteristics

Expanded Abstract and Analysis Prepared Especially for I.S.E. by David P. Butts, The University of Georgia.

Purpose

This study was designed to investigate how the science curriculum influences students' attitudes.

Rationale

To judge the value of a curriculum, a teacher needs to know what cognitive and affective outcomes are usually associated with that curriculum. Much study has been given to cognitive achievement outcomes but relatively sparse information is available on affective outcomes.

Research Design and Procedure

A one-shot post test design was used in which data from intact sixth grade classes were gathered on two dependent measures.

$$\begin{array}{cc} X_A & O_1 \\ X_B & O_2 \end{array}$$

where

X_A is "Process Approach instruction with 132 students

X_B is "Textbook Oriented" instruction with an undefined number of students

O_1 and O_2 include student preference on school subjects and a 21-item attitude scale.

Percentages of students' responses were tabulated for comparison.

Findings

Based on comparisons of percentages of students' responses from the two groups, the authors concluded that the "Process Approach" had a more positive influence on student attitudes than the "Textbook Oriented" approach.

ABSTRACTOR'S ANALYSIS

This brief report of an important area of study reads very well. The authors have provided a straightforward presentation of a researchable question, a design for its study, and conclusion. The introduction to the study could have been strengthened by coupling their questions to a rationale from previous research as to why it is reasonable to expect students' preference for science, or their attitudes toward science, to be influenced by the curriculum.

The design of the study offered an opportunity to explore possible relationships between important variables. However, the authors appear to believe that "exposure to the Process Approach" can be equated as a definable treatment or independent variable. Both a description of the independent variables and the estimates of validity that the two groups ("Process Approach" and "Textbook Oriented") were indeed different are missing from the report. While the authors do indeed select their sample to include an independent variable of "number of years exposure," the data analysis does not seem to have included alternatives to how "years in program" relates to attitude outcomes.

Two dependent variables are identified but credibility of the findings of the study is substantially reduced due to the absence of any reported validity or reliability estimates for either procedure.

The question of this study leads the reader to expect an exploratory or correlational study to identify possible relationships. The presentation of findings as description of percentages provides the reader with no answer to the question of the study. The authors seem to conclude that a higher percentage response (23 precent vs. 19 percent) or (79 percent vs. 49 percent) between two undefined independent variables means that one is proven more effective in influencing student attitudes than the other. While we might believe that to be true and have an adequate experience base to provide support for our belief, no evidence is presented in this report that can be used to influence our previously held conclusions.

INSTRUCTION

Holliday, William G. "The Effects of Verbal and Adjunct Pictorial-Verbal Information in Science Instruction." Journal of Research in Science Teaching, 12(1):77-83, 1975.

Descriptors--Biology, Educational Research, Instruction, *Learning Processes, *Measurement, *Pictorial Stimuli, Science Education, Secondary Education, *Secondary School Science, *Verbal Learning

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Robert H. Evans and Ronald D. Anderson, University of Colorado.

Purpose

The purpose of this study was to determine whether, on a non-pictorial criterion test, biology students using verbal text materials with supplementary line drawings would outperform students using only verbal texts.

Rationale

The author's analysis of the literature produces little support for the common belief that text material supplemented with pictures increases verbal learning. This conclusion is tempered when the various categories of "pictures" such as maps are included and when the criterion is pictorial rather than verbal. The author views most available research on the effectiveness of instructional pictures as either out of date, methodologically inadequate, or ungeneralizable.

The attention hypothesis and subsequent learning processes proposed by R. C. Anderson serve as the theoretical basis for this research. Specifically, these processes which the author relates to the experimental treatment materials are:

1. noticing the stimulus
2. encoding or processing the stimulus in a "meaningful" way
3. generating linkages between cues and responses.

Research Design and Procedure

Eighty randomly selected tenth-grade general biology students from two Calgary, Alberta high schools were randomly assigned to two instructional treatments. The seven classes from which the students were drawn generally were made up of students from the upper two-thirds of their grade level.

The first instructional treatment consisted of a 23-page verbal description of the effects of a plant growth hormone on stem and root growth. The other treatment used the same verbal description plus block line drawings having adjacent verbal descriptions of each drawing and, in some cases, verbal labels.

A verbal 30-item multiple choice test using 12 experimental situations based on the learning materials was designed as the criterion test. In an attempt to motivate students, both groups were told that their test performance was related to their ability to understand science material.

A One-Way Analysis of Variance (ANOVA) was used to test the difference between the treatment groups on the criterion measure.

Findings

The ANOVA yielded a significant difference at the .05 level in favor of the students who had the instructional materials with text plus pictures. The criterion test showed a Cronbach's Alpha of .86.

Interpretations

The author carefully concludes that this specific kind of picture and text combination can significantly facilitate one form of verbal comprehension. This finding is contrary to the generalizations of much research. He recommends that the theoretical foundations of this study be used as the basis for further research.

ABSTRACTOR'S ANALYSIS

This study provides a fresh and apparently useful theoretical basis for future research into multi-media system components in science. Specifically, the utilization of R. C. Anderson's learning process model is an important conceptual asset.

Even though the author provides a theoretical basis for the educational process being studied, the study is difficult to place within a context of related studies in this area. This difficulty is largely due to the lack of a conceptual framework within which the many studies in this area can be organized. While the author does a better job than most researchers of reviewing the past research, a void exists which is not filled by any of the past work cited or by the conceptual framework which he attempts to develop.

Few aspects of a research study are as likely to be given short shrift, yet play such an important role, as the review of previous research. Empirical research is a public and incremental process. The results are published where they are open to anyone interested. This public process is incremental in that the extant research becomes the basis for determining an appropriate focus for future research.

For past research to provide this foundation for future research, it must be synthesized into a meaningful conceptual framework. The subsequent research, in turn, should contribute to the knowledge organized within this conceptual framework.

The lack of a meaningful conceptual framework is a serious difficulty in dealing with the previous research in the area under consideration here. Without this needed synthesis (in contrast to a simple compilation) of past research results, it is difficult to determine if the study under consideration does indeed address a significant question or produce results that add to our understanding of the area.

The basic difficulty with the reviews cited is the failure to deal with this research in specific and detailed enough terms. For example, the

simple presence or absence of visual stimuli is too broad a question for a study to address with substantial hope of advancing knowledge in the field. The effectiveness of the visual stimuli may depend upon the purpose it serves (e.g., advance organizer vs. adjunct to the entire text; or cue vs. reinforcer), the nature of the content (e.g., age, stage of intellectual development, sex), or the criterion variable. These factors are not presented here as "the" list of items which should be studied, but simply to illustrate the type of factors which should be considered.

An effective review in this area should synthesize the past research in a manner which considers such factors and integrates the results into a conceptual framework with the potential of giving meaning to the extant and future research in the field. It should provide a means for determining which of the various factors have been important in the previous research and it thus should provide a basis for identifying promising future research. The absence of the desired conceptual framework for this study makes it difficult to decide if the researcher has addressed a significant question or made a significant contribution to knowledge.

In addition to the rationale for the study, a few other matters deserve comment in this review of Holliday's work. Considering the avilable diversity and complexity of pictorial instructional stimuli, as the author does, and associated issues such as textual placement, it is surprising that only two instructional formats are experimentally considered. A more comprehensive study, including for example a variety of picturial placements and formats would have been of more interest. Pictorial criterion questions, in addition to the verbal ones administered, might have revealed more clearly the unique learning contributions of the pictorial instructional stimuli. Again, an appropriate conceptual framework for the research in this area would provide a basis for choosing between the various alternatives vying for inclusion in the study.

Among other matters, there appears to be a lack of substantiation of the author's decision to place his pictures "in close proximity" to the related textual material. It is assumed, based on a report showing that students consider learning to be primarily a verbal activity, that they must be "led" to pictures. Placement would seem to be a highly appropriate variable to study within the author's theoretical model of learning in this area.

Holiday's report includes an example of one page of the pictorial/text material, and a sample verbal multiple choice test item about the sample. Based on this limited evidence, some concerns arise concerning the nature of the author-designed materials. While the study claims that the test is of a verbal nature, the given sample item specifically elicits a visual image of movement. Coincidentally, the associated pictorial/text material contains a prominent "arrow" which, if recalled, will answer the posed movement question. If this type of question is typical, the author-designed test, which is of an intended verbal nature, may be confounding the study outcomes by giving an advantage to students who had the pictorial material. However, it would be necessary to examine the other material employed in this study to resolve this matter.

In summary, it is agreed that future research in this area can benefit substantially by considering the theoretical requirements for learning outlined in this study. In addition, however, it would be most helpful to

have a thorough analysis of the extant research in the field with a detailed synthesis of past results. Such a synthesis would be of great benefit in delineating further studies to be conducted and interpreting the results of studies now in process.

48

Howe, Ann C. and David P. Butts. "The Effect of Instruction on the Acquisition of Conservation of Volume." Journal of Research in Science Teaching, 7(4):371-375, 1970.

Descriptors--*Conservation (Concept), *Concept Formation, *Elementary School Science, Learning, *Learning Readiness, Performance Tests, *Scientific Concepts

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Michael Szabo, Pennsylvania State University.

Purpose

The study by Howe and Butts served two purposes: (1) to determine the effects of exposure to the SAPA curriculum upon transition to formal operations (i.e., conservation of volume) of fourth and sixth grade students, and (2) to determine the effects of special instruction based upon a learning hierarchies scheme on the criterion of a Learning Hierarchies Test.

Rationale

The rationale for this study was that selected science instructional experiences should impact the levels of cognitive development of children as described by Piaget. The authors suggest the possibility that children who had received such curriculum materials would perform Piagetian tasks (e.g., indicative of formal operations) at an earlier age or in greater proportion than children not so treated. They also suggest that children whose intellectual development is "greater" but lack necessary information should perform tasks at a higher level after having been provided the information by the programs.

Research Design and Procedure

The study was a two group, pre- and posttest design which used a sample of 189 fourth and sixth grade students from two different schools. Some of the students had instruction in SAPA for at least fifteen months, while others had no such instruction. Assignment to SAPA was not conducted using principles of randomization; assignment to the special curriculum was randomized.

Prior to the experiment, SAPA students were compared with non-SAPA students on a volume concepts pretest based upon Piagetian tasks modified for group administration. The major independent variable was exposure to a special instructional curriculum devised to conform with a Learning Hierarchies Test developed by the authors. Students with and without SAPA experience were randomly assigned to either the instructional treatment group or to a control group.

The main criterion consisted of scores on a specially constructed Learning Hierarchies Test. In addition the correlation between age and performance on the tasks was investigated.

Findings

Prior to the actual experimental study, it was found that higher proportions of fourth grade SAPA students had successful scores on the volume concept test task #1 than non-SAPA students. This difference was not found for either fourth or sixth grade students on the second volume task. Using the criterion of Learning Hierarchies Test scores, the SAPA students out-scored the non-SAPA students at the fourth and sixth grade levels.

From the actual experiment, which compared students randomly assigned to instructional versus control groups within SAPA and non-SAPA groups nested within grade level, it was concluded that the instructional group out-scored the control group in three out of four situations which were tested using the Learning Hierarchies Test. In the fourth grade, the instructional group outscored the control group regardless of their SAPA background. In the sixth grade, however, there was no difference between the instructional and control groups who had the SAPA experience while the difference was significant for the non-SAPA group. The instructional program did not bring about any change in the proportion of children who could perform the volume conservation tasks.

Interpretations

The authors concluded that the ability to perform on criterion volume tasks is unrelated to either previous experience in the specified science program or to experience in the special instructional program. There was, however, a relationship between performance of the criterion volume tasks and age, grade level, and score on the Learning Hierarchies Test.

ABSTRACTOR'S ANALYSIS

The authors are to be praised for encouraging research which links practical curriculum matters with theoretical learning psychology models.

The reviewer's comments fall into issues of problem conceptualization and design. Can researchers expect that exposure to special curricula will indeed impact developmental levels? Flavell (1963) argues "yes" and "no" on the topic of speeding the acquisition of conservation through instructional practices. Flavell observes:

> Probably the most certain conclusion is that it can be a surprisingly difficult undertaking to manufacture Piagetian concepts in a laboratory. Almost all the training methods reported impress one as sound and reasonable and well-suited to the educative job at hand. And yet most of them have had remarkably little success in producing cognitive change . . . Further, there is more than a suspicion from present evidence that when one does succeed in inducing some behavioral change through this or that training procedure, it may not cut very deeply (p. 377).

An interpretation is that Piagetian structures are not artifacts of verbal confusion and misunderstandings (hence, not subject to amelioration through instruction), rather they are real and exert weight in the child's intellectual life.

Flavell's "yes" argument stems from Piaget's equilibrium model, which posits a sequence in which disequilibrium (antagonism between assimilation and accommodation) precedes acquisition of invariant functions. It has been argued that providing appropriate instruction to children in stages of disequalibrium will increase the likelihood that they will proceed to the next developmental stage "on schedule" without undue time delays. Smedslund's research (e.g., 1961a, 1961b) used a cognitive conflict strategy to stimulate the essential condition for the development of conservation, where previously there had been nonconservation, with encouraging results.

The above arguments lead to the expectations that exposure to the SAPA curriculum would not increase the numbers of subjects who could perform the criterion conservation task at an earlier age; however, it might increase the proportion of students performing the criterion conservation task if the SAPA curriculum:

1. fostered new cognitive conflict or attenuated existing cognitive conflict, and

2. provided information structures capable of adequately reducing the conflict, resulting in a return to equilibrium and conservation.

Since the results do not support the relation between SAPA experience and conservation of volume, the assumptions above may be questioned. Certainly the creators of SAPA do not claim to have designed the curriculum around cognitive conflict as it relates to Piagetian theory.

However, other questions related to methodology call for answers. First, subjects were not (and could not have been) randomly assigned to SAPA or non-SAPA, raising questions about developmental differences in the two groups. Developmental selection factors may have been in operation, as were non-developmental factors (e.g., the group differences in CTMM scores reported, but not the direction of the differences, by the authors).

Further, the only description of the non-SAPA group was that they ". . . had not had such instruction," leaving open the possibility of wide differences in experience and exposure to instruction as sound (developmentally) as SAPA.

Second, the validity of the measure of the conservation tasks may be questioned. Decisions about acquisition of conservation rest not only upon correctness of response but also on the soundness of the explanation provided. One can question whether the explanations were captured and used in the decision-making process in the group administered Test of Volume Concepts.

In the major experiment of the study, the treatment group registered significantly higher gains on the Learning Hierarchies Test but failed to increase the proportion of children who could perform the criterion conservation tasks.

Conceptually, one can argue that the most effective learning should occur when the "structure of the discipline" (in this case, a Gagnéian type learning hierarchy) is in high correspondence with the emerging cognitive structures (in this case, Piagetian volume conservation) of the subjects.

The reviewer hypothesizes that an adequate match between the discipline structure and the cognitive structure did not eventuate in this study.

The results suggest that there are real differences between models of intellectual development. An instructional approach based upon a content model of the development of intellect (Gagnè) was built and applied to a model of intellectual development characterized by rather invariant progression through a series of cognitive stages or levels. That the former did not impact one criterion variable from the latter model further supports the uniqueness and distinction between the two models.

The matching issue described above has more general implications for curriculum and instruction. Does the structure of any scientific discipline, as captured by subject matter experts, correspond in any meaningful fashion to the acquisition of knowledge by the uninitiated and inexperienced learner characterized by a pliable and emerging cognitive structure?

One solution path yet unexplored is curriculum research and development which rigorously and thoughtfully explores the macroscopic features of the discipline to be taught in light of a detailed analysis of the characteristics and features of the emerging cognitive structure of the neophyte learner. In short, let's have more of Howe and Butts research but at a more macroscopic and detailed level.

REFERENCES

Flavell, J. H. The Developmental Psychology of Jean Piaget. Princeton: D. Van Nostrand Co., 1963.

Howe, A. C. and D. P. Butts. "The Effect of Instruction on the Acquisition of Conservation of Volume." Journal of Research in Science Teaching, 7:371-375, 1970.

Smedslund, J. "The Acquisition of Conservation of Substance and Weight in Children. V. Practice in Conflict Situations Without External Reinforcement." Scandinavian Journal of Psychology, 2:156-160, 1961.

Smedslund, J. "The Acquisition of Conservation of Substance and Weight in Children. VI. Practice on Continuous Versus Discontinuous Material in Conflict-Situations Without External Reinforcement." Scandinavian Journal of Psychology, 2:203-210, 1961.

Humphreys, Donald W. and Ronald D. Townsend. "The Effects of Teacher- and Student-Selected Activities on the Self-Image and Achievement of High School Biology Students." Science Education, 58(3):295-301, 1974.

Descriptors--*Achievement, *Biology; Educational Research, Individualized Curriculum, *Individualized Instruction, *Secondary School Science, Secondary School Students, Science Education, *Self Concept.

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Gene Gennaro, University of Minnesota.

Purpose

Two types of individualized instruction were investigated to determine what effect freedom for the student to develop or select his/her learning experiences had upon (1) his/her academic achievement, (2) how well he/she perceived he/she could achieve academically, and (3) how much time he/she required to gain concept competency.

Rationale

The authors suggest (1) that John Carroll's model of mastery provides a way to satisfactorily master science concepts; (2) that individualized instruction (in this case, self-paced) should provide the student with the required time for developing factual information and experiences into meaningful concepts and principles; (3) that perhaps the ultimate goal of learning experiences should be encouragement of the student's independence in learning, permitting him to build broad principles and using them to solve specific problems; that (4) Carroll's model of mastery learning provides the opportunity for a greater percentage of students to be successful. However, the authors also suggest that students who take more time to achieve mastery may experience a feeling of failure.

Research Design and Procedure

The sample consisted of two BSCS Green Version Biology classes from Owatonna High School, Owatonna, Minnesota. The students in both classes were responsible for learning the same biological concepts using individualized instruction. In one class, the teacher provided the procedures to solve science problems, which led to concept understanding [teacher-structured experience group (TSE)]. The students in the other class were given only the science problems and were required to develop or select a set of experiences which led toward concept understanding [student-structured experience group (SSE)]. Using means on pre-test scores in comparing both classes using the IOWA Tests of Educational Development (Natural Science), IQ, Q-sort of a single adjective, and pre-test scores on the BSCS Comprehensive Final as indicators, and after performing a t-test on each of the four tests, the authors state that the students in both classes were from the same population. In both

groups concept understanding was tested periodically by using concept tests which required them to use the learned concept to solve an unfamiliar problem. If the concept test was not passed, the students were given additional experiences which were prescribed (TSE group) or developed new experiences from a number of resources (SSE group). The BSCS comprehensive final examination was used to measure differences in academic achievement between the TSE and SSE groups after 140 days of classroom studies. Self-image of academic achievement was measured by a single adjective Q-sort technique at the beginning of the experiment and after 47, 92, and 120 days of classroom activity. The cards were sorted by the student so that the sorted adjectives represented how the student perceived his/her ability to achieve.

Findings

Using a t-test, it was determined that the TSE group scored significantly higher on the post-test of the BSCS Comprehensive Final than did the SSE group. For the four measures of self-image in academic achievement taken at the beginning and after 47, 92, and 120 days of classroom activity, the fluctuation of group means was greater in the SSE group than in the TSE group. There was also a drop "to a significant degree" in the means of scores measuring SSE students' self-image between the first and second Q-sorts even though it was reported that the overall self-image of achievement was not significantly different. However, these scores did vary significantly among students within each group.

Significantly more concepts were developed by the TSE group during the experimental period than by the SSE group as shown by the means of time in days required for concept understanding (18.6 days vs. 23.2 days respectively). A negative correlation was found between the self-image of achievement and the time needed to achieve the concepts as shown by lower self-concept scores of SSE students. Similarly, a negative correlation existed between the actual academic achievement and the time factor.

When the achievement scores were plotted, the authors report, it became evident that the TSE scores had a more central grouping while the greatest extremes existed in the SSE group.

Interpretations

The differences in means on the self-image scores between the SSE group and the TSE group on the second test reflect, the authors of the articles suggest, the early level of frustration demonstrated by the SSE group, which was not as prevalent in the TSE group. The authors suggest that the difference in frustration level could have resulted from a greater change in learning pattern required of the SSE group than was required of the TSE group. During the study, independence of learning required gradual development by most students. The TSE group, the authors report, was initially significantly superior in their ability to achieve academically and required significantly less time to pass the "individual concept tests.

The difference between the mean times in student days for concept understanding (23.2 for SSE group; 18.6 for TSE group) can be accounted for, the authors suggest, by the greater time required by the students of the SSE group to construct and select their learning experiences, while the TSE group needed no time to do this.

Although the TSE groups' academic performance was superior to that of the SSE group, there is an indication that some students in the SSE group were able to develop an independent learning behavior, as indicated by superior schievement scores. This implication, the authors suggest, is inconclusive and warrants further research.

ABSTRACTOR'S ANALYSIS

The study is getting at some important learning considerations. Do many or most students need the organization of the teacher to provide maximum learning? Is it useful to give students the choice of materials and learning modes to achieve mastery of learning concepts if in the process of taking longer to achieve understanding of the chosen concepts, the students' academic self-image deteriorates?

There is some indication in this study to suggest that academic self-image decreases at first but then increases to the point where, at the end of 120 days, there is no significance in academic self-image between students in the SSE group compared to students in the TSE group. The authors suggest that there is an early level of frustration demonstrated by the SSE group, which is not as prevalent in the TSE group and that the difference in frustration level could have resulted from a greater change in learning pattern required of the SSE group than was required of the TSE group. There is also some indication that certain groups of students profit by structuring their own learning as shown by the achievement of some students (three) in the SSE group who outscored any student in the TSE group.

A negative correlation was shown between the self-image of achievement and the time needed to achieve concept understanding. Could this be due to the fact that students who take longer to achieve mastery include students who are not motivated academically? There was, also, a negative correlation between the academic achievement and the time factor. Could this be due to the same reason?

The authors, by using certain measurable characteristics such as IQ scores, natural science achievement scores, academic self-image scores and pre-test scores on the test to be used as a final measurement of cognitive learning (the BSCS Comprehensive Final), attempted to show that the TSE group and SSE group were from the same population. Critics can point out that, without randomization, other factors such as persistence to a task, the ability to organize one's learning activity, the preferred mode of learning for students in the study - just to mention a few factors - may be equally important in affecting learning rate, amount of material learned, and academic self-image. It would be useful to repeat the study, randomizing the students, in the hope that these and other factors are equally distributed.

The authors state in the "conclusions" that the TSE group was initially significantly superior in their ability to achieve academically and required significantly less time to pass the individual concept tests." I was unable to verify this, unless this is inferred by the fact that the TSE group required significantly less time to pass the individual concept tests.

Throughout the report, 29 students are mentioned as being involved in the SSE sample; however, only 28 are accounted for on the graph which shows the frequency of achievement scores. It would be well to mention what happened to the one student.

The mean of the pre-test for the SSE students on the BSCS Comprehensive Final (20.13) is above that of the score on the post-test for the same group (19.69). Even the TSE group did not gain much from the pre- to post-test using the BSCS Comprehensive Final (20.46, 22.57 respectively). The use of a standardized test is laudable; however, one can question the validity of the BSCS Comprehensive Final for what was taught to both groups of students.

It would be well to include in a new study a third group of students enrolled in a course of biology, in which mastery was not the learning model used, and where defining the level of performance before proceeding was not the mode used. This would allow us to compare the achievement and academic self-image with students in a mastery mode. It may be that requiring students to master concepts as shown by scores on periodic tests before moving on may act detrimentally to both achievement and academic self-image. It might be that this has a particular influence on attitude and thence on achievement. More research needs to be carried out in this area.

Mastery may be applicable for basic understandings of some selected material in science, but may not be appropriate for some of the kind of exploratory and experimental work that students do in a science class. Those activities which get students to utilize process aspects of science may not be an essential part of a mastery mode of teaching but may, however, be more motivating and may result in better understanding of science process and content.

In certain areas of learning, it may be appropriate to use a mastery model such as the kind Carroll suggests in learning the skills necessary in areas such as arithmetic, reading, foreign language vocabulary, etc., or even in the mastery of certain concepts in <u>all</u> disciplines. However, the learning which cannot be labeled skill or concept development such as enrichment, experiential learning or process learning in certain aspects of science where mastery is not necessarily sought, requires a different mode of presentation and may in fact result in longer term gains by making the learner more excited about what is being learned, and hence affect learning outcomes.

Kempa, R. F. and J. E. Ward. "The Effect of Different Modes of Task Orientation on Observational Attainment in Practical Chemistry." Journal of Research in Science Teaching, 12(1):69-76, 1975.
Descriptors--*Chemistry, Educational Research, *Instruction, *Observational Learning, Problem Solving, *Performance Factors, Science Education, Secondary Education, Secondary School Science, *Task Performance

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Vincent N. Lunetta, The University of Iowa.

Purpose

To assess the effects of three different modes of instruction (task orientation) on student observational skills in laboratory work in chemistry.

Rationale

The authors base the need for the study upon the fact that laboratory work in contemporary science courses is central and should involve the student in problem-solving and in the process of scientific inquiry. Hence, it is important to examine the attributes and to identify the conditions under which the educational effectiveness of laboratory work may be optimized. The authors describe four phases of practical work in science (planning and design, manipulation, observation, and interpretation) and examine the observation phase in this study. They cite four laboratory courses, each of which embodies one of the three idealized modes of laboratory task orientations defined in the study. The paper does not refer the reader to other relevant research studies.

Research Design and Procedure

The three modes of laboratory task orientation selected for study were:

(a) an open-ended approach in which observational tasks had to be accomplished in the absence of any form of cueing;

(b) a method of partial direction in which students received cueing to some but not all observations to be made;

(c) a check-list approach in which students were required to carry out observational tasks with reference to a comprehensive schedule listing all possible observations.

A chemistry observation test was prepared that consisted of ten test-tube reactions in four perceptual areas: ". . . color changes; changes involving the formation or disappearance of solids; changes involving the liberation of gases; and temperature changes resulting from the evolution or absorption of heat during a reaction." The ten reactions were selected through a series of pretests to ensure that results would not depend on students' manipulative ability and that students would be unfamiliar with the actual chemical system. Students recorded their observations on

report sheets that had a format consistent with one of the three modes of task orientation, but all students were provided with standardized sets of equipment and had "ready access to all solutions and materials required." Sufficient time was allowed for all students to complete the experiments and reports.

> Completion times for the chemistry obseration tests including the recording of observations, ranged from 70 to 90 minutes . . . . Scores derived . . . were treated as criterion-referenced measures since the test aimed at establishing absolute levels of observational competence achieved as the result of the three treatment modes . . . .

Two other tests were administered within one week of the observation test to provide supplemental information. In an assessment of color vision, all students in the study correctly described and differentiated between colors of solutions on fourteen color slides. To assess the effects of academic ability, a short multiple-choice chemistry achievement test was administered.

The sample selected for the study consisted of 140 fourth-year "O"-level chemistry students in three schools. The subjects in each school were randomly assigned to the three task orientation groups.

Mean scores, standard deviations, and sample size were reported for each of the three task orientation treatments and schools. Also the percent mean observational error rates were reported for each treatment and school. Two categories of error rates were reported for each of the treatments: errors of omission occurring when the subject failed to perceive clearly observable changes, and illusory errors arising when observations were reported that were not present in the system being observed. Mean scores for each of the three treatments were also reported for students assigned to three achievement groups on the basis of their performance on the chemistry achievement test. Finally, mean observational error rates were reported as a function of the complexity of the particular observational task.

Analyses of variance were conducted to determine the effects of the three task orientation treatments and to determine the effects of the independent variables.

Findings

1. In all schools students in the check-list groups scored highest and students in the partial-direction group performed least well on the observational tasks.

2. Observational attainment was not significantly influenced by the school or by a treatment mode-school interaction.

3. No significant relationship was observed between students' cognitive abilities in chemistry and observational attainments.

4. The mean percentage for errors of omission was much higher for students in the open-ended mode than in the check-list mode. The partial direction mode achieved an even higher rate for errors of omission; further analysis of this group revealed a major difference between cued and noncued observations with the error rate especially high on noncued observations.

5. The mean percentage for illusory errors was very much higher for students in the check-list mode than for students in either of the other two treatment modes.

6. As the observational task became more complex, students in all treatment groups failed to perceive more of the stimuli that were present.

Interpretations

1. Observational skills in laboratory chemistry "are not primarily 'intellectual' in nature."

2. The relatively low error rate of students in the check-list group can be attributed to the cueing that mode provides, particularly in complex observational situations.

3. A differential effect of cueing on directed and nondirected observational tasks can be observed. "The performance level on the latter appears substantially impaired as the result of the specific cueing given to other observations."

4. The conclusion that the check-list mode is superior based on an analysis of the omissional errors must be made with reservations. The high observational attainment for that group may well be a consequence of its extensive illusory error rate. (There was a direct relationship between illusory error and mean observational attainment.)

In final analysis, the conclusions reported by the investigators were:

1. Observational attainment in laboratory chemistry is significantly influenced by instructional conditions.

2. The check-list mode of task orientation appears to be most effective in producing high observational achievement, but induces a high illusory error rate which is almost absent in the other two modes.

3. The partial direction mode results in a high success rate on those observations for which specific cueing has been provided, but inhibits the effective perception of other, noncued stimuli.

4. The relative merits of the open-ended and the check-list modes cannot be fully assessed since no studies have as yet been conducted to examine how students interpret "sets of observations which are either incomplete (because of omission errors) or include information which is illusory in origin."

ABSTRACTOR'S ANALYSIS

This study is a particularly important one for research in science education because it examines a topic that is unique to science laboratory teaching. The design of the study provides a model for objective investigations on other topics in science education. Currently, very inadequate attention has been given to research into the "practical mode" of laboratory teaching, and the mechanisms set up in the study may provide insights to other researchers who will examine the development of laboratory or practical skills.

Although research on laboratory skills is very inadequate, the investigators provide no reference to any related investigations; neither do they comment on the dearth of relevant studies. The failure to bridge the gap between the present study and relevant prior research is a weakness of the written paper. The authors also do not address sufficiently the implications of their investigation for further research study. While they have pointed out one or two questions that require further research, they have not discussed the broad need for related research, nor have they discussed the implications of techniques developed in their own research for such study. Implications of the study for teaching practice are implicit in the paper, but they are not elaborated at length.

In general, the implications of this interesting study were not discussed sufficiently. Specific suggestions could have been directed to the relationship between specific instructional techniques and the goals of instruction, and a variety of significant questions for further investigations could have been elaborated. Also, while the authors were relatively cautious in the generalizing from their data, they should have cited the need for replication studies, in their conclusions section, due to the limitations of the sample and of the laboratory tasks. The sample consisted of a relatively small, unique group of students who were examined on one set of experiments with one set of observational criteria and with one set of instructional tasks. While the study has broad implications, the generalizations that can be drawn from it at this time are considerably narrower than the title of the paper implies.

The analysis of data and the attention to detail within that data is impressive in the written paper. The authors have been through in their research and, generally, have reported details with care. Nevertheless, the reader may raise some of the following questions that were not answered in the written report.

Why was the study limited to cognitive measures? What are the effects of task orientation on the attitudes of students?

How does prior experience affect observational attainment? The pretest was probably effective in eliminating topics that had previously been experienced in chemistry, but the entry level of students in observational skills was not examined, and the sample may well have had some bias in the development of these skills relative to a broader population of students.

How do different modes of task orientation affect observational attainment over long periods of time? The study examined a 70 to 90 minute sequence of tests and observations. What would occur over a period of days or months?

What are the effects of intellectual development, or intelligence, or achievement generally? The "chemistry achievement test" was brief and highly specialized and did not assess intellect or competence broadly. Thus, data are insufficient for the statement that observational abilities "are not primarily 'intellectual' in nature" (p. 73).

What are the effects of other related skills like reading ability? Would an orally administered test produce different results?

What would be learned by an orally administered protocol that would assess observational attainment?

Are hierarchical relationships present in observational skills? Observational task complexity is defined in the study as the number of correct observations possible in a given experiment, and the study does not investigate the question of hierarchical relationships.

The study also does not examine the relationship between the four major phases of practical work in science delineated in the paper's introduction. "The four phases are not only sequential, they are also hierarchical . . . success at one stage determines the success achievable at the next stage" (p. 69). This assumption of relationship is not used or examined in the ensuing report.

Sparks, Phillip D. and Laraine M. Unbehaun. "Achievement of Audio-Tutorial and Conventional Biology Students, A Comparative Study." BioScience, 21(12):574-576, 1971.

Descriptors--*Autoinstructional Programs, *Biology, *College Science, *Educational Research, *Instruction, Research, Science Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by William S. LaShier, Jr., University of Kansas.

Purpose

The purpose of this study was to evaluate the achievement of college biology students using an audio-tutorial (A-T) program by comparing their performance with students using a conventional biology lecture-laboratory program. The four null hypotheses included:

1. Initially the students of the experimental group were not significantly different from the control group as measured by scores on the Natural Science portion of the American College Test (ACT).

2. The post-achievement of students in the Audio-Tutorial (experimental) group was not significantly different from that of students in the control group as measured by the Total Biology Test.

3. The scores on the ACT pretest and the scores on the Total Biology Test are uncorrelated for both the control and the experimental groups.

4. The percentage of the students initially enrolled who are included in the study is the same for both the control and the experimental groups.

Rationale

This study, published in 1971, was apparently one of only a few A-T comparative studies using a control vs. experimental group format with groups numbering over a hundred subjects. The unstated assumption of the study seemed to be that retention of factual information was the most important dependent variable to measure. The experimental treatment known as audio-tutorial instruction was nearly identical to the procedures popularized by Postlethwait, et al. (1969).

Research Design and Procedure

A posttest control group design was used to contrast the achievement of students using an audio-tutorial program (A-T) with that of students in a conventional biology lecture-laboratory course. The 190 students in the A-T group met in a weekly general assembly for a unit introduction and for evaluation purposes. Students were free to spend as much time as desired

in independent study involving audio-taped lectures and laboratory materials. Small sessions of approximately 30 students met once a week with an instructor to review difficult concepts. In contrast, the 180 students in the conventional or control group attended three one-hour lectures and one two-hour laboratory per week. The subject matter presented to both groups was the same.

The length of the instruction of the college biology A-T and conventionally taught groups was not stated. The original sizes of the A-T group and the control group were eventually reduced to 153 and 143, respectively, because of withdrawals and other reasons.

The Natural Science portion of the American College Test was used to test the null hypothesis of no initial differences in the two groups. The Total Biology Test was administered to both groups. This test was actually an accumulation of 274 test items which had been administered in part periodically throughout the course. of the 274 items, 165 were later categorized into nine subtests corresponding to topics taught in the course. The Total Biology Test was principally a measure of the retention of factual information with approximately 10 percent of the items at the application level.

The mean achievement scores on the 10 tests for the two groups were analyzed using the z-test statistic for evidence of significant difference at the .05 level. A similar analysis was used with the pretest data to determine if there were any initial differences between the two groups. In each of these statistical computations, all the student scores were used and then just the freshman student scores were used since the original control group contained a significant number of nonfreshmen.

Findings

The test of the first hypothesis indicated that there was no significant difference between the students in the A-T section and in the conventionally taught section when the students entered the course. Hypothesis four confirmed that the percentage drops in student enrollment in the experimental and control groups were not significantly different. Hypothesis three pointed to significant Pearson product-moment correlation of 0.45 and 0.46 between the pretest and the Total Biology Test for the A-T and conventionally-taught students, respectively.

The analysis of hypothesis two indicated significant differences in achievement scores on the Total Biology Test and on three subtests involving the topics of chemistry of life, plant reproduction, and ecology-evolution. In all cases of significant differences, the A-T group scored higher than the control group. These differences held true even when only the freshman scores were used.

Interpretations

The authors said that the results of the study indicate that students using an A-T format achieved more than those using a conventional format. In addition, 90 percent of the students in the A-T section indicated that they would select sections of other general education courses that were

taught using the A-T approach in preference to a conventional method. The authors pointed out that the reader should expect future contradictory evidence in studies of A-T systems because of such things as the variability of the format, and variability of course objectives.

ABSTRACTOR'S ANALYSIS

The biology audio-tutorial programs for college students appear to be further distinguished by the self-directed and self-paced features of instruction. The A-T programs should not be confused with an individualized program which, in addition to the self-paced feature, makes the instructional material available in more than one presentation format (Rowe and Deture, 1975).

The authors of the present study were quite correct in predicting that future studies contrasting achievement gains in A-T groups would furnish contradictory evidence. Rowe and Deture reviewed four 1974 studies on A-T College Biology vs. Traditional programs. The scorecard showed two studies with no content achievement differences, and one study favoring the A-T approach, and one study favoring the traditional approach.

The research design used in the present study did not provide for a truly random assignment of students to each of the groups. Instead, the students in the experimental group signed up for the A-T section and the control group signed up for the conventional lecture laboratory session. The finding that a significantly high number of non-freshmen signed up for the control group suggests that an element of bias may have entered into the composition of the two groups. The subsequent reanalysis of the data using only freshmen did not necessarily remove the question of the random nature of both groups.

The research design also provided for only one instructor for the conventionally taught group of students, while the A-T group could come into contact with several instructors. Since in many traditional programs students attend a large lecture and then are placed with different teaching assistants for the laboratory, it would seem that the use of one person to handle all the conventional instruction could be advantageous if the individual coordinated the two phases of instruction.

The authors of the present study indicated that both groups were presented the same subject matter. However, a more specific explanation of the support aids was not given. It would have been interesting, for instance, to know whether both groups were given behavioral objectives, similar handouts, and similar laboratory activities. Perhaps of even more value would have been some figures on the relative amount of time each group spent working on biology tasks. The outcome of the study was generally limited to factual recall information. It would have been interesting to have had comparative information on science process skill development and attitudinal measures.

The authors reported that 90 percent of the students enrolled in the A-T section indicated they would select a similar A-T format for other courses in the future. However, no mention was made of whether the control group student would prefer to take future courses in the lecture-lab format. It would seem possible since the control group students apparently had the option of signing up for the A-T section but did not.

The audio-tutorial format is especially suited for aptitude-treatment interaction studies. These types of studies attempt to determine whether certain students with particular cognitive learning styles, attitudes or aptitudes achieve better with a particular treatment such as audio-tutorial instruction. Grobe (1973), in a comparison of A-T with a conventional biology program, grouped students in three aptitude levels with scores on the CEEB being the dependent variable. With the knowledge of aptitude and treatment the author was only able to account for 9.5 percent of the variance, suggesting the need to search for additional variables.

A study of two different strategies for sequencing five biology units within an audio-tutorial course was conducted by Gunter (1974). The SBUS sequence represented a logical development of concepts. The TDUS sequence represented a logical development of concepts. The TDUS sequence provided special sequence of topics arranged from units the student knew best to the least understood units as determined by student pretest scores. The results of the study indicated no significant differences between the two groups in terms of achievement or attitude toward the course. The interesting point, however, was that the TDUS students completed the work in half the time it took the SBUS group.

Future researchers may wish to pursue the time spent on task variables as an important predictor of achievement in Audio-Tutorial instruction. In another phase of the study by Gunter, the audio-tutorial students with high grade point averages elected to spend more time in completing the unit and in addition scored higher on the achievement posttests than did students with low grade point averages. Rowe and Deture (1975) in their review of audio-tutorial research commented on the widespread concern of educators with student procrastination. The evidence provided by Gunter suggests that additional external guidance may be necessary to keep students with low grade point averages on task longer and more consistently.

Another area of possible research was suggested by Meleca's (1973) description of the Bio-Learning Center at Ohio State University. In addition to a choice of multidimensional learning materials keyed to behavioral objectives, and access to small-group recitation sections, the program provides a broadening experience. That is, instructional television programs are provided which deal with the short range goals of helping the students. Programs are also provided dealing with the broader implications of biology and research. The movement toward more individualized instruction in the A-T format will provide new impetus for measuring higher cognitive and affective objectives.

An audio-tutorial approach to biology instruction in the high school was reported by Nordland (1975). No significant differences in achievement were observed between the A-T and conventional biology class. The authors also reported that standardized measures such as the SAT reading and the STEP science tests were less predictive of learner achievement when the audio-tutorial methods were used. This finding of lower correlations suggests that the A-T format might provide a more appropriate instructional alternative for students deficient in certain skills such as reading.

REFERENCES

1. Grobe, Cary H. "A Regression Approach to Evaluating Instructional Programs in Science." Journal of Research in Science Teaching, 10(1):55-62, 1973.

2. Gunter, Alfred V. "The Effects of Different Sequences of Instructional Units and Experiences Within Instructional Units on the Achievement and Attitudes of College General Biology Students." Dissertation Abstracts, 34(11):7066A, 1974.

3. Herron, J. Dudley, Harold H. Jaus, Van Neie, Thom Luce, and Terry O'Heron. A Summary of Research in Science Education - 1974. A Supplement Volume of Science Education. New York: John Wiley and Sons, 1976 (pp. 17-22).

4. Meleca, C. Benjamin. "Ohio State University Biology-Learning Center Uses the A-T Method." American Biology Teacher, 35(4):192-197, 1973.

5. Nordland, Floyd H., Jane B. Kahles, Stephen Randak, and Thomas Watts. "An Analysis of the Effectiveness of Audio-Tutorial Instruction: Measured by Student Achievement and Predicted by Standardized Measures." School Science and Mathematics, 59:277-284, 1974.

6. Postlethwait, S. N., J. Novak, and H. T. Murray, Jr. The Audio-Tutorial Approach to Learning. Minneapolis: Burgess Publishing Company, 1969.

7. Rowe, Mary Budd and Linda DeTure. A Summary of Research in Science Education - 1973. New York: John Wiley and Sons, 1976 (pp. 16-22).